A discontinuous-skeletal method for advection-diffusion-reaction on general meshes

Daniele Antonio Di Pietro, Jerome Droniou, Alexandre Ern

A discontinuous-skeletal method for advection-diffusion-reaction on general meshes

Daniele A. Di Pietro*[1], Jérôme Droniou†[2], and Alexandre Ern‡[3]

[1]University of Montpellier 2, I3M, 34057 Montpellier CEDEX 5, France
[2]School of Mathematical Sciences, Monash University, Clayton, Victoria 3800, Australia
[3]University Paris-Est, CERMICS (ENPC), 6–8 avenue Blaise Pascal, 77455 Marne-la-Vallée CEDEX 2, France

November 1, 2014

Abstract

We design and analyze an approximation method for advection-diffusion-reaction equations where the (generalized) degrees of freedom are polynomials of order $k \geqslant 0$ at mesh faces. The method hinges on local discrete reconstruction operators for the diffusive and advective derivatives and a weak enforcement of boundary conditions. Fairly general meshes with polytopal and nonmatching cells are supported. Arbitrary polynomial orders can be considered, including the case $k = 0$ which is closely related to Mimetic Finite Difference/Mixed-Hybrid Finite Volume methods. The error analysis covers the full range of Péclet numbers, including the delicate case of local degeneracy where diffusion vanishes on a strict subset of the domain. Computational costs remain moderate since the use of face unknowns leads to a compact stencil with reduced communications. Numerical results are presented.
2010 Mathematics Subject Classification: 65N30, 65N08,

1 Introduction

The goal of the present work is to design and analyze an approximation method for advection-diffusion-reaction equations where the (generalized) degrees of freedom (DOFs) are polynomials of order $k \geqslant 0$ at mesh faces. Since such faces constitute the mesh skeleton, and since DOFs can be chosen independently at each face, we use the terminology discontinuous-skeletal method. The proposed method offers various assets: (i) Fairly general meshes, with polytopal and nonmatching cells, are supported; (ii) Arbitrary polynomial orders, including the case $k = 0$, can be considered; (iii) The error analysis covers the full range of Péclet numbers; (iv) Computational costs remain moderate since skeletal DOFs lead to a compact stencil with reduced communications.

Approximation methods using face-based DOFs have been investigated recently for advection-diffusion equations on meshes composed of standard elements. In [6], Cockburn et al. devise and numerically investigate a Hybridizable Discontinuous Galerkin (HDG) method for the diffusion-dominated regime based on a mixed formulation where an approximation for the total advective-diffusive flux is sought. In [5], Chen and Cockburn carry out a convergence analysis for a variable degree HDG method on semimatching nonconforming simplicial meshes, and investigate the impact of mesh nonconformity on the supercloseness of the potential. The formulation differs from [6] in that the flux variable now approximates the diffusive component only. In [23], Wang and Ye

*daniele.di-pietro@univ-montp2.fr, corresponding author
†jerome.droniou@monash.edu
‡ern@cermics.enpc.fr

analyze a Weak Galerkin method for advection-diffusion-reaction on triangular meshes, which appears to be mainly tailored to the diffusion-dominated case. Turning to general polyhedral meshes, we cite, in particular, the work of Beirão da Veiga, Droniou, and Manzini [2] on Hybrid Mimetic Mixed (HMM) methods (which encompass, see [15], three families of numerical schemes for elliptic equations: the Mimetic Finite Difference method [3], the Mixed Finite Volume method [14], and the Hybrid Finite Volume – or SUSHI – method [18]). Although the analysis focuses on the diffusion-dominated case, we show in this work that a suitable tweaking of the scheme so as to include weakly enforced boundary conditions allows one to treat the advection-dominated case as well.

The starting point for the present discontinuous-skeletal method is the Hybrid-High Order (HHO) method designed in [9, 11] for purely diffusive and linear elasticity problems. The key ideas in [9,11] are as follows: (i) In each mesh cell, a local potential reconstruction of order $(k+1)$ is devised from polynomials of order k in the cell and on its faces (cell- and face-based DOFs); (ii) A local bilinear form is built using a Galerkin form based on the gradient of the local potential reconstruction plus a stabilization form which preserves the improved order of the reconstruction; this leads to energy-error estimates of order $(k+1)$ and L^2-potential estimates of order $(k+2)$ if elliptic regularity holds; (iii) The global discrete problem is assembled cellwise, and cell-based DOFs are eliminated by static condensation, so that only the face-based DOFs remain.

The extension to advection-diffusion-reaction equations entails several new ideas: (i) We devise a local reconstruction of the advective derivative from cell- and face-based DOFs using an integration by parts formula; (ii) Stability for the advective contribution is ensured by terms that penalize the difference between cell- and face-based DOFs at faces, and which therefore do not preclude the possibility of performing static condensation and do not enlarge the stencil; as in [2], the stability terms are formulated in a rather general form so as to include various approaches used in the literature, e.g., upwind, locally θ-upwind, and Scharfetter–Gummel schemes; (iii) Boundary conditions are enforced weakly so as to achieve robustness in the full range of Péclet numbers.

An additional novel feature of the present work is that our analysis also includes the case of locally degenerate advection-diffusion-reaction equations, where the diffusion coefficient vanishes on a (strict) subset of the computational domain. We emphasize that such problems are particularly delicate since the exact solution can jump at the diffusive/nondiffusive interface separating zero and nonzero regions for the diffusion coefficient. The literature on locally degenerate advection-diffusion-reaction problems is relatively scarce. The coupling of parabolic-hyperbolic systems in one space dimension is considered by Quarteroni and Gastaldi [19]; see also [17]. In both cases, ad hoc techniques are proposed based on removing suitable terms at the diffusive/nondiffusive interface. In [10], a discontinuous Galerkin (dG) method is designed and analyzed, where the use of weighted averages allows one to automatically handle the possibility of jumps in the exact solution. A dG method handling the degenerate case is also considered numerically by Houston, Schwab, and Süli [21]. In the present setting, a perhaps surprising result at first sight is that an approximation method hinging on face-based DOFs can capture well a discontinuous solution. This result is achieved owing to a tailored design of the stabilization terms ensuring that the interface unknown approximates well the exact solution from the diffusive side.

The material is organized as follows. In Section 2, we present the model problem. In Section 3, we introduce the discrete setting. In Section 4, we define the local bilinear forms and introduce the novel ideas to discretize the advective terms. In Section 5, we build the discrete problem, introduce several norms for the analysis, and state our main results on stability and error estimates. The dependence on the physical parameters is tracked in the error estimate so as to capture the variation in convergence order between the diffusive and the advective regimes. We also study the link with the HMM methods of [2] in the case $k = 0$. Such a link was already noticed in [11] for HHO methods in the purely diffusive case. Finally, numerical results on standard and general polygonal meshes are presented to assess the sharpness of the error estimate both in the uniformly vanishing diffusion and locally degenerate case. Finally, in Section 6, we prove our main results.

2 Model problem

Let $\Omega \subset \mathbb{R}^d$, $d \geqslant 1$, be an open bounded connected polytope of boundary $\partial\Omega$ and unit outer normal $\boldsymbol{n}$. We denote by $\nu : \Omega \to \mathbb{R}^+$ the diffusion coefficient, which we assume to be piecewise constant on a partition $P_\Omega := \{\Omega_i\}_{1\leqslant i\leqslant N_\Omega}$ of Ω into polytopes and such that $\nu \geqslant \underline{\nu} \geqslant 0$ almost everywhere in Ω. The case of locally heterogeneous anisotropic diffusion can be considered as well using the ideas in [8]. For the advective velocity $\boldsymbol{\beta} : \Omega \to \mathbb{R}^d$, we assume the regularity $\boldsymbol{\beta} \in \mathrm{Lip}(\Omega)^d$, and for the reaction coefficient $\mu : \Omega \to \mathbb{R}$, we assume and that μ is bounded from below by a real number $\mu_0 > 0$. For simplicity, we work under the assumption $\boldsymbol{\nabla}\cdot\boldsymbol{\beta} \equiv 0$; the case $\boldsymbol{\nabla}\cdot\boldsymbol{\beta} \not\equiv 0$ can be treated similarly, provided $\mu + \frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{\beta} \geqslant \mu_0 > 0$. We introduce the following sets (cf. Figure 3 below for an illustration):

$$\Gamma_{\nu,\boldsymbol{\beta}} := \{\boldsymbol{x} \in \partial\Omega \mid \nu > 0 \text{ or } \boldsymbol{\beta}\cdot\boldsymbol{n} < 0\}, \tag{1a}$$

$$\mathcal{I}^{\pm}_{\nu,\boldsymbol{\beta}} := \{\boldsymbol{x} \in \mathcal{I}_\nu \mid \pm(\boldsymbol{\beta}\cdot\boldsymbol{n}_I)(\boldsymbol{x}) > 0\}, \tag{1b}$$

where $\mathcal{I}_\nu$ is the diffusive/nondiffusive interface and $\boldsymbol{n}_I$ is the unit normal to $\mathcal{I}_\nu$ pointing out of the diffusive region. More precisely, $\mathcal{I}_\nu$ is the set of points in Ω located at an interface between two distinct subdomains Ω_i and Ω_j of P_Ω such that $\nu_{|\Omega_i} > \nu_{|\Omega_j} = 0$. We assume that

$$(\boldsymbol{\beta}\cdot\boldsymbol{n}_I)(\boldsymbol{x}) \neq 0 \text{ for a.e. } \boldsymbol{x} \in \mathcal{I}_\nu.$$

For given source term $f \in L^2(\Omega)$ and boundary datum $g \in L^2(\Gamma_{\nu,\boldsymbol{\beta}})$, the continuous problem reads

$$\boldsymbol{\nabla}\cdot(-\nu\boldsymbol{\nabla}u + \boldsymbol{\beta}u) + \mu u = f \qquad \text{in } \Omega\backslash\mathcal{I}_\nu, \tag{2a}$$

$$[\![-\nu\boldsymbol{\nabla}u + \boldsymbol{\beta}u]\!]\cdot\boldsymbol{n}_I = 0 \qquad \text{on } \mathcal{I}_\nu, \tag{2b}$$

$$[\![u]\!] = 0 \qquad \text{on } \mathcal{I}^+_{\nu,\boldsymbol{\beta}}, \tag{2c}$$

$$u = g \qquad \text{on } \Gamma_{\nu,\boldsymbol{\beta}}, \tag{2d}$$

where $[\![\cdot]\!]$ denotes the jump across $\mathcal{I}_\nu$ (the sign is irrelevant). Notice that the boundary condition is enforced at portions of the boundary touching a diffusive region or a nondiffusive region provided the advective field flows into the domain. A weak formulation for (2) has been analyzed in [10]. In the non-degenerate case $\underline{\nu} > 0$, $\Gamma_{\nu,\boldsymbol{\beta}} = \partial\Omega$ and the usual weak formulation in the space $H^1_0(\Omega)$ holds, cf., e.g., [7, Section 4.6.1].

3 Discrete setting

This section presents the discrete setting: admissible mesh sequences, analysis tools on such meshes, DOFs, reduction maps, and reconstruction operators.

3.1 Assumptions on the mesh

Denote by $\mathcal{H} \subset \mathbb{R}^+_*$ a countable set of meshsizes having 0 as its unique accumulation point. Following [7, Chapter 4], we consider h-refined mesh sequences $(\mathcal{T}_h)_{h\in\mathcal{H}}$ where, for all $h \in \mathcal{H}$, $\mathcal{T}_h$ is a finite collection of nonempty disjoint open polyhedral elements T such that $\overline{\Omega} = \bigcup_{T\in\mathcal{T}_h} \overline{T}$ and $h = \max_{T\in\mathcal{T}_h} h_T$ with h_T standing for the diameter of the element T. A face F is defined as a hyperplanar closed connected subset of $\overline{\Omega}$ with positive $(d-1)$-dimensional Hausdorff measure and such that (i) either there exist $T_1(F), T_2(F) \in \mathcal{T}_h$ such that $F \subset \partial T_1(F) \cap \partial T_2(F)$ and F is called an interface or (ii) there exists $T(F) \in \mathcal{T}_h$ such that $F \subset \partial T(F) \cap \partial\Omega$ and F is called a boundary face. In what follows, the dependence on F of $T_1(F)$ and $T_2(F)$ (when F is an interface) and of $T(F)$ (when F is a boundary face) is omitted when no ambiguity can arise. Interfaces are collected in the set $\mathcal{F}^{\mathrm{i}}_h$, boundary faces in $\mathcal{F}^{\mathrm{b}}_h$, and we let $\mathcal{F}_h := \mathcal{F}^{\mathrm{i}}_h \cup \mathcal{F}^{\mathrm{b}}_h$. The diameter of a face $F \in \mathcal{F}_h$ is denoted by h_F. For all $T \in \mathcal{T}_h$, $\mathcal{F}_T := \{F \in \mathcal{F}_h \mid F \subset \partial T\}$ denotes the set of faces contained

in ∂T (with ∂T denoting the boundary of T) and, for all $F \in \mathcal{F}_T$, $\boldsymbol{n}_{TF}$ is the unit normal to F pointing out of T. Symmetrically, for all $F \in \mathcal{F}_h$, we let $\mathcal{T}_F := \{T \in \mathcal{T}_h \mid F \subset \partial T\}$ the set of elements having F as a face. For each interface $F \in \mathcal{F}_h^{\mathrm{i}}$, we fix an orientation as follows: we select a fixed ordering for the elements $T_1, T_2 \in \mathcal{T}_h$ such that $F \subset \partial T_1 \cap \partial T_2$ and we let $\boldsymbol{n}_F := \boldsymbol{n}_{T_1,F}$. For a boundary face, we simply take $\boldsymbol{n}_F = \boldsymbol{n}$, the outward unit normal to Ω. In what follows, we denote by $|\cdot|_l$ the l-dimensional Hausdorff measure.

Our analysis hinges on the following two assumptions on the mesh sequence.

Assumption 1 (Admissible mesh sequence)**.** *For all $h \in \mathcal{H}$, $\mathcal{T}_h$ admits a matching simplicial submesh $\mathfrak{T}_h$ and there exists a real number $\varrho > 0$ independent of h such that, for all $h \in \mathcal{H}$, (i) for all simplex $S \in \mathfrak{T}_h$ of diameter h_S and inradius r_S, $\varrho h_S \leqslant r_S$ and (ii) for all $T \in \mathcal{T}_h$, and all $S \in \mathfrak{T}_h$ such that $S \subset T$, $\varrho h_T \leqslant h_S$.*

Assumption 2 (Compatible mesh sequence)**.** *(i) Any mesh cell belongs to one and only one subdomain Ω_i of the partition P_Ω; (ii) Any mesh face having an intersection with the interface $\mathcal{I}_{\nu,\boldsymbol{\beta}}$ (of positive $(d-1)$-dimensional Hausdorff measure) is included in one of the two sets $\mathcal{I}_{\nu,\boldsymbol{\beta}}^{\pm}$; (iii) In any mesh face such that the diffusion coefficient vanishes on both of its sides, the normal component of $\boldsymbol{\beta}$ is nonzero in a subset of positive measure.*

The simplicial submesh in Assumption 1 is just a theoretical tool used to prove the results in Section 3.2, and it is not used in the actual construction of the discretization method. Furthermore, a straightforward consequence of Assumption 2(i) is that ν is piecewise constant on $\mathcal{T}_h$. Assumption 2(ii) is important in the error analysis so that the face unknowns on $\mathcal{I}_{\nu,\boldsymbol{\beta}}$ capture the exact solution from the diffusive side. Assumption 2(iii) can be avoided by adding some crosswind diffusion to the stabilization of the advective-reactive bilinear form in the spirit of a Lax–Friedrichs flux, so that the difference between cell- and face-based DOFs is always penalized on faces included in the nondiffusive region.

3.2 Analysis tools

We recall some results that hold uniformly in h on admissible mesh sequences. In what follows, for $X \subset \overline{\Omega}$, we respectively denote by $(\cdot,\cdot)_X$ and $\|\cdot\|_X$ the standard inner product and norm in $L^2(X)$, with the convention that the subscript is omitted whenever $X = \Omega$. The same notation is used in the vector-valued case $L^2(X)^d$. According to [7, Lemma 1.42], for all $h \in \mathcal{H}$, all $T \in \mathcal{T}_h$, and all $F \in \mathcal{F}_T$, h_F is comparable to h_T in the sense that

$$\varrho^2 h_T \leqslant h_F \leqslant h_T. \tag{3}$$

Moreover, [7, Lemma 1.41] shows that there exists an integer N_∂ depending on ϱ such that

$$\forall h \in \mathcal{H}, \qquad \max_{T\in\mathcal{T}_h} \operatorname{card}(\mathcal{F}_T) \leqslant N_\partial. \tag{4}$$

Let $l \geqslant 0$ be a nonnegative integer. For an n-dimensional subset X of $\overline{\Omega}$ ($n \leqslant d$), $\mathbb{P}_n^l(X)$ is spanned by the restriction to X of n-variate polynomials of total degree $\leqslant l$. Then, there exists a real number C_{tr} depending on ϱ and l, but independent of h, such that the following discrete trace inequality holds for all $T \in \mathcal{T}_h$ and $F \in \mathcal{F}_T$, cf. [7, Lemmas 1.46]:

$$\|v\|_F \leqslant C_{\mathrm{tr}} h_F^{-1/2} \|v\|_T \qquad \forall v \in \mathbb{P}_d^l(T). \tag{5}$$

Furthermore, the following inverse inequality holds for all $T \in \mathcal{T}_h$ with C_{inv} again depending on ϱ and l, but independent of h, cf. [7, Lemma 1.44],

$$\|\boldsymbol{\nabla} v\|_T \leqslant C_{\mathrm{inv}} h_T^{-1} \|v\|_T \qquad \forall v \in \mathbb{P}_d^l(T). \tag{6}$$

Moreover, using [7, Lemma 1.40] together with the results of [16], one can prove that there exists a real number C_{app} depending on ϱ and l, but independent of h, such that, for all $T \in \mathcal{T}_h$, denoting by π_T^l the L^2-orthogonal projector on $\mathbb{P}_d^l(T)$, the following holds: For all $s \in \{1, \dots, l+1\}$ and all $v \in H^s(T)$,

$$|v - \pi_T^l v|_{H^m(T)} + h_T^{1/2} |v - \pi_T^l v|_{H^m(\partial T)} \leqslant C_{\mathrm{app}} h_T^{s-m} |v|_{H^s(T)} \qquad \forall m \in \{0, \dots, s-1\}. \tag{7}$$

3.3 Degrees of freedom, interpolation, and reconstruction

Let a polynomial degree $k \geqslant 0$ be fixed. For all $T \in \mathcal{T}_h$, the local space of DOFs is

$$\underline{\mathsf{U}}_T^k := \mathbb{P}_d^k(T) \times \left\{ \bigtimes_{F \in \mathcal{F}_T} \mathbb{P}_{d-1}^k(F) \right\},$$

and we use the notation $\underline{\mathsf{v}}_T = (\mathsf{v}_T, (\mathsf{v}_F)_{F \in \mathcal{F}_T})$ for a generic element $\underline{\mathsf{v}}_T \in \underline{\mathsf{U}}_T^k$. We define the local interpolation operator $\underline{\mathsf{I}}_T^k : H^1(T) \to \underline{\mathsf{U}}_T^k$ such that, for all $v \in H^1(T)$,

$$\underline{\mathsf{I}}_T^k v := \left(\pi_T^k v, (\pi_F^k v)_{F \in \mathcal{F}_T} \right),$$

where π_F^k denotes the L^2-orthogonal projector on $\mathbb{P}_{d-1}^k(F)$. Following [11], for all $T \in \mathcal{T}_h$, we define the local potential reconstruction operator $p_T^k : \underline{\mathsf{U}}_T^k \to \mathbb{P}_d^{k+1}(T)$ such that, for all $\underline{\mathsf{v}}_T := (\mathsf{v}_T, (\mathsf{v}_F)_{F \in \mathcal{F}_T}) \in \underline{\mathsf{U}}_T^k$,

$$\begin{aligned} (\boldsymbol{\nabla} p_T^k \underline{\mathsf{v}}_T, \boldsymbol{\nabla} w)_T &= (\boldsymbol{\nabla} \mathsf{v}_T, \boldsymbol{\nabla} w)_T + \sum_{F \in \mathcal{F}_T} (\mathsf{v}_F - \mathsf{v}_T, \boldsymbol{\nabla} w \cdot \boldsymbol{n}_{TF})_F \qquad \forall w \in \mathbb{P}_d^{k+1}(T), \\ \int_T p_T^k \underline{\mathsf{v}}_T &= \int_T \mathsf{v}_T. \end{aligned} \tag{8}$$

The discrete Neumann problem (8) is well-posed. The following result has been proved in [11, Lemma 3].

Lemma 1 (Approximation properties for $p_T^k \underline{\mathsf{I}}_T^k$)**.** *There exists a real number $C > 0$, depending on ϱ and k, but independent of h_T, such that, for all $v \in H^{k+2}(T)$,*

$$\begin{aligned} &\|v - p_T^k \underline{\mathsf{I}}_T^k v\|_T + h_T^{1/2} \|v - p_T^k \underline{\mathsf{I}}_T^k v\|_{\partial T} \\ &\qquad + h_T \|\boldsymbol{\nabla}(v - p_T^k \underline{\mathsf{I}}_T^k v)\|_T + h_T^{3/2} \|\boldsymbol{\nabla}(v - p_T^k \underline{\mathsf{I}}_T^k v)\|_{\partial T} \leqslant C h_T^{k+2} \|v\|_{H^{k+2}(T)}. \end{aligned} \tag{9}$$

4 Local bilinear forms

In this section we define the local bilinear forms. These forms are expressed in terms of local DOFs and are instrumental to derive the discrete problem in Section 5.

4.1 Diffusion

To discretize the diffusion term in (2), we introduce, for all $T \in \mathcal{T}_h$, the bilinear form $a_{\nu,T}$ on $\underline{\mathsf{U}}_T^k \times \underline{\mathsf{U}}_T^k$ such that

$$a_{\nu,T}(\underline{\mathsf{w}}_T, \underline{\mathsf{v}}_T) := (\nu_T \boldsymbol{\nabla} p_T^k \underline{\mathsf{w}}_T, \boldsymbol{\nabla} p_T^k \underline{\mathsf{v}}_T)_T + s_{\nu,T}(\underline{\mathsf{w}}_T, \underline{\mathsf{v}}_T), \tag{10}$$

with stabilization bilinear form $s_{\nu,T}$ on $\underline{\mathsf{U}}_T^k \times \underline{\mathsf{U}}_T^k$ such that

$$s_{\nu,T}(\underline{\mathsf{w}}_T, \underline{\mathsf{v}}_T) := \sum_{F \in \mathcal{F}_T} \frac{\nu_T}{h_F} (\pi_F^k(\mathsf{w}_F - P_T^k \underline{\mathsf{w}}_T), \pi_F^k(\mathsf{v}_F - P_T^k \underline{\mathsf{v}}_T))_F. \tag{11}$$

In (11), the potential reconstruction $P_T^k : \underline{\mathsf{U}}_T^k \to \mathbb{P}_d^{k+1}(T)$ is such that, for all $\underline{\mathsf{v}}_T \in \underline{\mathsf{U}}_T^k$,

$$P_T^k \underline{\mathsf{v}}_T := \mathsf{v}_T + (p_T^k \underline{\mathsf{v}}_T - \pi_T^k p_T^k \underline{\mathsf{v}}_T),$$

where the second term can be interpreted as a high-order correction of v_T.

4.2 Advection-reaction

For all $T \in \mathcal{T}_h$, we introduce the discrete advective derivative $G^k_{\boldsymbol{\beta},T} : \underline{\mathsf{U}}^k_T \to \mathbb{P}^k_d(T)$ such that, for all $\underline{\mathsf{v}}_T \in \underline{\mathsf{U}}^k_T$ and all $w \in \mathbb{P}^k_d(T)$,

$$(G^k_{\boldsymbol{\beta},T}\underline{\mathsf{v}}_T, w)_T = -(\mathsf{v}_T, \boldsymbol{\beta}\cdot\boldsymbol{\nabla} w)_T + \sum_{F\in\mathcal{F}_T} ((\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF})\mathsf{v}_F, w)_F \tag{12a}$$

$$= (\boldsymbol{\beta}\cdot\boldsymbol{\nabla}\mathsf{v}_T, w)_T + \sum_{F\in\mathcal{F}_T} ((\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF})(\mathsf{v}_F - \mathsf{v}_T), w)_F, \tag{12b}$$

where we have integrated by parts the first term in the right-hand side and used $\boldsymbol{\nabla}\cdot\boldsymbol{\beta} \equiv 0$ to pass from (12a) to (12b). We introduce, for all $T \in \mathcal{T}_h$, the local bilinear form $a_{\boldsymbol{\beta},\mu,T}$ on $\underline{\mathsf{U}}^k_T \times \underline{\mathsf{U}}^k_T$ such that

$$a_{\boldsymbol{\beta},\mu,T}(\underline{\mathsf{w}}_T, \underline{\mathsf{v}}_T) := -(\mathsf{w}_T, G^k_{\boldsymbol{\beta},T}\underline{\mathsf{v}}_T)_T + (\mu\mathsf{w}_T, \mathsf{v}_T)_T + s^-_{\boldsymbol{\beta},T}(\underline{\mathsf{w}}_T, \underline{\mathsf{v}}_T).$$

The local stabilization bilinear forms $s^\pm_{\boldsymbol{\beta},T}$ on $\underline{\mathsf{U}}^k_T \times \underline{\mathsf{U}}^k_T$ are such that

$$s^\pm_{\boldsymbol{\beta},T}(\underline{\mathsf{w}}_T, \underline{\mathsf{v}}_T) := \sum_{F\in\mathcal{F}_T} (\tfrac{\nu_F}{h_F} A^\pm(\mathrm{Pe}_{TF})(\mathsf{w}_F - \mathsf{w}_T), \mathsf{v}_F - \mathsf{v}_T)_F.$$

Here, $\nu_F := \min_{T\in\mathcal{T}_F} \nu_T$ is the lowest diffusion coefficient from the (one or) two cells sharing F, and the local (oriented) Péclet number Pe_{TF} is defined if $\nu_F > 0$ by

$$\mathrm{Pe}_{TF} = h_F \frac{\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF}}{\nu_F}, \tag{13}$$

while we use (14) below if $\nu_F = 0$. Since, for all $F \in \mathcal{F}^{\mathrm{b}}_h$, there is a unique $T \in \mathcal{T}_h$ such that $F \subset \partial T$, we simply write Pe_F instead of Pe_{TF} in this case. Notice that the local Péclet number Pe_{TF} is a function $F \to \mathbb{R}$.

The functions $A^\pm : \mathbb{R} \to \mathbb{R}$ are such that $A^\pm(s) = \frac{1}{2}(|A|(s) \pm s)$ for all $s \in \mathbb{R}$, and the function $|A| : \mathbb{R} \to \mathbb{R}$ is assumed to satisfy the following design conditions:

- (A1) $|A|$ is a Lipschitz-continuous function such that $|A|(0) = 0$ and, for all $s \in \mathbb{R}$, $|A|(s) \geqslant 0$ and $|A|(-s) = |A|(s)$;
- (A2) there exists $\underline{a} \geqslant 0$ such that $|A|(s) \geqslant \underline{a}|s|$ for any $|s| \geqslant 1$;
- (A3) If $\underline{\nu} = 0$, $\lim_{s\to+\infty} \frac{|A|(s)}{s} = 1$ (and, consistently with (A1)), $\lim_{s\to-\infty} \frac{|A|(s)}{s} = -1$. Coherently, for all $T \in \mathcal{T}_h$ and all $F \in \mathcal{F}_T$ such that $\nu_F = 0$, we set

$$\tfrac{\nu_F}{h_F} A^\pm(\mathrm{Pe}_{TF}) := \lim_{\nu\to 0^+} \left(\tfrac{\nu}{h_F} A^\pm\left(\tfrac{h_F}{\nu}\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF}\right)\right) = (\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF})^\pm, \tag{14}$$

  where, for a real number s, we have denoted $s^\pm := \frac{1}{2}(|s| \pm s)$.

As already pointed out in [2, 4, 13], using the generic functions $A^\pm$ in the definition of the advective terms allows for a unified treatment of several classical discretizations. The centered scheme corresponds to $|A|(s) = 0$, which fails to satisfy (A2)-(A3). Instead, Properties (A1)–(A3) are fulfilled by the following methods:

- *Upwind scheme*: $|A|(s) = |s|$ (so that $A^\pm(s) = s^\pm$ and $\frac{\nu_F}{h_F} A^\pm(\mathrm{Pe}_{TF}) = (\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF})^\pm$).
- *Locally upwinded θ-scheme*: $|A|(s) = (1-\theta(s))|s|$, where $\theta \in C^1_c(-1,1)$, $0 \leqslant \theta \leqslant 1$ and $\theta \equiv 1$ on $[-1/2, 1/2]$, corresponding to the centered scheme if $s \in [-1/2, 1/2]$ (dominating diffusion) and the upwind scheme if $s \geqslant 1$ (dominating advection).
- *Scharfetter–Gummel scheme*: $|A|(s) = 2\left(\frac{s}{2}\coth(\frac{s}{2}) - 1\right)$.

The advantage of the locally upwinded θ-scheme and the Scharfetter–Gummel scheme over the upwind scheme is that they behave as the centered scheme, and thus introduce less artificial diffusion, when s is close to zero (dominating diffusion).

Remark 2 (Assumption (A2)). *This assumption implies that $|A^{\pm}(s)| \leqslant \overline{a}|A(s)|$ for all $|s| \geqslant 1$ with $\overline{a} = \frac{1}{2}(1 + \underline{a}^{-1})$. Furthermore, the threshold $|s| \geqslant 1$ is arbitrary. If it is changed into $|s| \geqslant \underline{b}$ for some fixed $\underline{b} \geqslant 1$, then the only modification in the error estimate* (29) *below is to change the term* $\min(1, \mathrm{Pe}_T)$ *into* $\min(\underline{b}, \mathrm{Pe}_T)$.

Remark 3 (Assumption (A3)). *Assumption (A3) is only required in the locally degenerate case where the diffusion coefficient vanishes in one part of the domain.*

5 Discrete problem and main results

We build in this section the discretization of (2) using the local bilinear forms of Section 4. A key point is the weak enforcement of boundary conditions to achieve robustness with respect to the Péclet number.

5.1 Discrete bilinear forms

Local DOFs are collected in the following global space obtained by patching interface values:

$$\underline{\mathsf{U}}_h^k := \left\{ \bigtimes_{T\in\mathcal{T}_h} \mathbb{P}_d^k(T) \right\} \times \left\{ \bigtimes_{F\in\mathcal{F}_h} \mathbb{P}_{d-1}^k(F) \right\}.$$

We use the notation $\underline{\mathsf{v}}_h = ((\mathsf{v}_T)_{T\in\mathcal{T}_h}, (\mathsf{v}_F)_{F\in\mathcal{F}_h})$ for a generic element $\underline{\mathsf{v}}_h \in \underline{\mathsf{U}}_h^k$ and, for all $T \in \mathcal{T}_h$, it is understood that $\underline{\mathsf{v}}_T$ denotes the restriction of $\underline{\mathsf{v}}_h$ to $\underline{\mathsf{U}}_T^k$.

Denoting by $\varsigma > 0$ a user-dependent boundary penalty parameter, we define the global diffusion bilinear form $a_{\nu,h}$ on $\underline{\mathsf{U}}_h^k \times \underline{\mathsf{U}}_h^k$ such that

$$a_{\nu,h}(\underline{\mathsf{w}}_h, \underline{\mathsf{v}}_h) := \sum_{T\in\mathcal{T}_h} a_{\nu,T}(\underline{\mathsf{w}}_T, \underline{\mathsf{v}}_T) + \sum_{F\in\mathcal{F}_h^{\mathrm{b}}} \left\{ -(\nu_F \boldsymbol{\nabla} p_{T(F)}^k \underline{\mathsf{w}}_T \cdot \boldsymbol{n}_{TF}, \mathsf{v}_F)_F + \frac{\varsigma \nu_F}{h_F} (\mathsf{w}_F, \mathsf{v}_F)_F \right\}, \quad (15)$$

and the global advection-reaction bilinear form $a_{\boldsymbol{\beta},\mu,h}$ such that

$$a_{\boldsymbol{\beta},\mu,h}(\underline{\mathsf{w}}_h, \underline{\mathsf{v}}_h) := \sum_{T\in\mathcal{T}_h} a_{\boldsymbol{\beta},\mu,T}(\underline{\mathsf{w}}_T, \underline{\mathsf{v}}_T) + \sum_{F\in\mathcal{F}_h^{\mathrm{b}}} (\tfrac{\nu_F}{h_F} A^+(\mathrm{Pe}_F) \mathsf{w}_F, \mathsf{v}_F)_F. \quad (16)$$

The rightmost terms in (15) and (16) are responsible for the weak enforcement of the boundary condition on $\Gamma_{\nu,\boldsymbol{\beta}}$. Finally, we set

$$a_h(\underline{\mathsf{w}}_h, \underline{\mathsf{v}}_h) := a_{\nu,h}(\underline{\mathsf{w}}_h, \underline{\mathsf{v}}_h) + a_{\boldsymbol{\beta},\mu,h}(\underline{\mathsf{w}}_h, \underline{\mathsf{v}}_h), \quad (17)$$

and we define the linear form l_h on $\underline{\mathsf{U}}_h^k$ such that

$$l_h(\underline{\mathsf{v}}_h) := \sum_{T\in\mathcal{T}_h} (f, \mathsf{v}_T)_T + \sum_{F\in\mathcal{F}_h^{\mathrm{b}}} \left\{ (\tfrac{\nu_F}{h_F} A^-(\mathrm{Pe}_F) g, \mathsf{v}_F)_F + \frac{\varsigma \nu_F}{h_F} (g, \mathsf{v}_F)_F \right\}. \quad (18)$$

The discrete problem reads: Find $\underline{\mathsf{u}}_h \in \underline{\mathsf{U}}_h^k$ such that, for all $\underline{\mathsf{v}}_h \in \underline{\mathsf{U}}_h^k$,

$$a_h(\underline{\mathsf{u}}_h, \underline{\mathsf{v}}_h) = l_h(\underline{\mathsf{v}}_h). \quad (19)$$

Remark 4 (Variations of $a_{\nu,h}$). *A symmetric expression of $a_{\nu,h}$ is obtained by adding the term $-\sum_{F\in\mathcal{F}_h^{\mathrm{b}}} (\mathsf{w}_F, \nu_F \boldsymbol{\nabla} p_{T(F)}^k \underline{\mathsf{v}}_T \cdot \boldsymbol{n}_{TF})_F$, to the right-hand side of* (15)*, and correspondingly the term $-\sum_{F\in\mathcal{F}_h^{\mathrm{b}}} (g, \nu_F \boldsymbol{\nabla} p_{T(F)}^k \underline{\mathsf{v}}_T \cdot \boldsymbol{n}_{TF})_F$ to the right-hand side of* (18)*. This variation is not further pursued here since the problem* (2) *is itself nonsymmetric.*

5.2 Discrete norms and stability

The analysis of the discrete problem (19) involves several norms. For the sake of easy reference, their definitions are gathered here, as well as some related stability properties. The energy-like diffusion (semi)norm is defined on $\underline{\mathsf{U}}_h^k$ by

$$\begin{aligned}&\|\underline{\mathsf{v}}_h\|_{\nu,h}^2 := \sum_{T\in\mathcal{T}_h} \|\underline{\mathsf{v}}_T\|_{\nu,T}^2 + |\underline{\mathsf{v}}_h|_{\nu,\partial\Omega}^2 \quad \text{with:}\\ &\|\underline{\mathsf{v}}_T\|_{\nu,T}^2 := a_{\nu,T}(\underline{\mathsf{v}}_T,\underline{\mathsf{v}}_T) \quad \text{and} \quad |\underline{\mathsf{v}}_h|_{\nu,\partial\Omega}^2 := \sum_{F\in\mathcal{F}_h^{\mathrm{b}}} \frac{\nu_F}{h_F}\|\mathsf{v}_F\|_F^2,\end{aligned} \tag{20}$$

and owing to [11, Lemma 3.1], we observe that there is $\eta>0$, depending only on ϱ, d, and k, such that, for all $\underline{\mathsf{v}}_T\in\underline{\mathsf{U}}_T^k$,

$$\eta\|\underline{\mathsf{v}}_T\|_{\nu,T}^2 \leqslant \nu_T\|\boldsymbol{\nabla}\mathsf{v}_T\|_T^2 + \sum_{F\in\mathcal{F}_T}\frac{\nu_T}{h_T}\|\mathsf{v}_F-\mathsf{v}_T\|_F^2 \leqslant \eta^{-1}\|\underline{\mathsf{v}}_T\|_{\nu,T}^2. \tag{21}$$

The advection-reaction (semi)norm is defined on $\underline{\mathsf{U}}_h^k$ by

$$\begin{aligned}&\|\underline{\mathsf{v}}_h\|_{\boldsymbol{\beta},\mu,h}^2 := \sum_{T\in\mathcal{T}_h} \|\underline{\mathsf{v}}_T\|_{\boldsymbol{\beta},\mu,T}^2 + |\underline{\mathsf{v}}_h|_{\boldsymbol{\beta},\partial\Omega}^2 \quad \text{with:}\\ &\|\underline{\mathsf{v}}_T\|_{\boldsymbol{\beta},\mu,T}^2 := \frac12\sum_{F\in\mathcal{F}_T} \|\big[\tfrac{\nu_F}{h_F}|A|(\mathrm{Pe}_{TF})\big]^{1/2}(\mathsf{v}_F-\mathsf{v}_T)\|_F^2 + \tau_{\mathrm{ref},T}^{-1}\|\mathsf{v}_T\|_T^2 \quad \text{and}\\ &|\underline{\mathsf{v}}_h|_{\boldsymbol{\beta},\partial\Omega}^2 := \frac12\sum_{F\in\mathcal{F}_h^{\mathrm{b}}} \|\big[\tfrac{\nu_F}{h_F}|A|(\mathrm{Pe}_F)\big]^{1/2}\mathsf{v}_F\|_F^2.\end{aligned} \tag{22}$$

Following [7, Chapter 2], the reference time $\tau_{\mathrm{ref},T}$ and velocity $\beta_{\mathrm{ref},T}$ are defined by

$$L_{\boldsymbol{\beta},T} := \max_{1\leqslant i\leqslant d}\|\boldsymbol{\nabla}\beta_i\|_{L^\infty(T)^d}, \quad \tau_{\mathrm{ref},T} := \{\max(\|\mu\|_{L^\infty(T)}, L_{\boldsymbol{\beta},T})\}^{-1}, \quad \beta_{\mathrm{ref},T} := \|\boldsymbol{\beta}\|_{L^\infty(T)^d}, \tag{23}$$

(recall that $\boldsymbol{\beta}\in\mathrm{Lip}(\Omega)^d$ implies $\boldsymbol{\beta}\in W^{1,\infty}(\Omega)^d$). Finally, we define two advection-diffusion-reaction norms on $\underline{\mathsf{U}}_h^k$ as follows:

$$\|\underline{\mathsf{v}}_h\|_{\flat,h}^2 := \|\underline{\mathsf{v}}_h\|_{\nu,h}^2 + \|\underline{\mathsf{v}}_h\|_{\boldsymbol{\beta},\mu,h}^2 \quad \text{and} \quad \|\underline{\mathsf{v}}_h\|_{\sharp,h}^2 := \|\underline{\mathsf{v}}_h\|_{\flat,h}^2 + \sum_{T\in\mathcal{T}_h} h_T\beta_{\mathrm{ref},T}^{-1}\|G_{\boldsymbol{\beta},T}^k\underline{\mathsf{v}}_T\|_T^2, \tag{24}$$

where the summand is taken only if $\beta_{\mathrm{ref},T}\neq 0$. The error estimate stated in Theorem 9 below uses the $\|\cdot\|_{\sharp,h}$-norm, and therefore delivers information on the advective derivative of the error, which is important in the advection-dominated regime. The $\|\cdot\|_{\flat,h}$-norm is, on the other hand, the natural coercivity norm for the bilinear form a_h, and is used as an intermediary step in the error analysis. The coercivity norm is sufficient for the error analysis in the diffusion-dominated regime.

Remark 5 (Norms)**.** *Owing to Assumption 2(iii), we infer that $\nu_F\neq 0$ or $\boldsymbol{\beta}\cdot\boldsymbol{n}_F\not\equiv 0$ (on a subset with positive measure). Hence, $\|\cdot\|_{\flat,h}$ and $\|\cdot\|_{\sharp,h}$ are norms on $\underline{\mathsf{U}}_h^k$. Indeed, if $\nu_F\neq 0$, then by* (21) *the diffusive norm controls the term $\mathsf{v}_F-\mathsf{v}_T$ and, if $\nu_F=0$, owing to* (14)*, we obtain $\frac{\nu_F}{h_F}|A|(\mathrm{Pe}_{TF}) = |\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF}|\not\equiv 0$ and the advective norm controls $\mathsf{v}_F-\mathsf{v}_T$.*

Our first important result concerns stability. The proof is postponed to Section 6.1.

Lemma 6 (Stability of a_h)**.** *Assume $\varsigma\geqslant 1+\frac{C_{\mathrm{tr}}^2N_\partial}{2}$ and (A1)–(A3). Then, for all $\underline{\mathsf{v}}_h\in\underline{\mathsf{U}}_h^k$, the following holds:*

$$\zeta\|\underline{\mathsf{v}}_h\|_{\flat,h}^2 \leqslant a_h(\underline{\mathsf{v}}_h,\underline{\mathsf{v}}_h), \tag{25}$$

with $\zeta := \min_{T\in\mathcal{T}_h}(\frac12,\tau_{\mathrm{ref},T}\mu_0)>0$. Assume additionally that, for all $T\in\mathcal{T}_h$,

$$h_TL_{\boldsymbol{\beta},T}\leqslant\beta_{\mathrm{ref},T} \quad \text{and} \quad h_T\mu_0\leqslant\beta_{\mathrm{ref},T}, \tag{26}$$

where $L_{\boldsymbol{\beta},T}$, $\beta_{\mathrm{ref},T}$, and $\tau_{\mathrm{ref},T}$ are defined by (23)*. Then, there exists a real number $\gamma > 0$, independent of h, ν, $\boldsymbol{\beta}$, and μ, such that, for all $\underline{\mathsf{w}}_h \in \underline{\mathsf{U}}_h^k$,*

$$\gamma\zeta\|\underline{\mathsf{w}}_h\|_{\sharp,h} \leqslant \sup_{\underline{\mathsf{v}}_h \in \underline{\mathsf{U}}_h^k \setminus \{0\}} \frac{a_h(\underline{\mathsf{w}}_h, \underline{\mathsf{v}}_h)}{\|\underline{\mathsf{v}}_h\|_{\sharp,h}}. \tag{27}$$

Remark 7 (Threshold for ς)**.** *The dependency on C_{tr} of the threshold on ς introduced in Lemma 6 can be removed by considering a lifting-based penalty term such as the one discussed in [7, Section 5.3.2] and originally introduced by Bassi, Rebay et al. [1] in the context of discontinuous Galerkin methods. Furthermore, the strict minimal threshold in Lemma 6 is $\varsigma > \frac{C_{\mathrm{tr}}^2 N_\partial}{4}$.*

Remark 8 (Assumption (26))**.** *The first inequality in* (26) *stipulates that the meshsize resolves the spatial variations of the advective velocity $\boldsymbol{\beta}$. The quantity $\mathrm{Da}_T := h_T \mu_0 \beta_{\mathrm{ref},T}^{-1}$ is a local Damköhler number relating the reactive and advective time scales. The second inequality in* (26) *assumes that $\mathrm{Da}_T \leqslant 1$ for all $T \in \mathcal{T}_h$, meaning that we are not concerned with the reaction-dominated regime. We could of course state a stability result without* (26)*, but the dependency on the various constants would be somewhat more intricate.*

5.3 Error estimate

For all $F \in \mathcal{F}_h$, we denote by $T_\nu(F)$ one element of $\mathcal{T}_F$ such that $T_\nu(F) \in \arg\max_{T \in \mathcal{T}_F} \nu_T$ (such an element may not be unique when F is an interface). Consider now an interface $F \in \mathcal{F}_h^{\mathrm{i}}$ such that $F \subset \mathcal{I}_{\nu,\boldsymbol{\beta}}^-$. Since the exact solution can jump on F, we have to deal with a possibly two-valued trace for the exact solution. It turns out that, in this case, the face unknown captures the trace from the diffusive side, i.e., from the unique element $T_\nu(F) \in \mathcal{T}_F$ such that $\nu_{|T_\nu(F)} > 0$. We therefore define the global interpolation operator $\underline{\mathsf{I}}_h^k : H^1(\Omega \backslash \mathcal{I}_{\nu,\boldsymbol{\beta}}) \to \underline{\mathsf{U}}_h^k$ is such that, for all $v \in H^1(\Omega \backslash \mathcal{I}_{\nu,\boldsymbol{\beta}})$:

$$\underline{\mathsf{I}}_h^k v := \left((\pi_T^k v)_{T \in \mathcal{T}_h}, (\pi_F^k [v_{|T_\nu(F)}])_{F \in \mathcal{F}_h} \right). \tag{28}$$

Our main result is the following estimate on the discrete approximation error $(\underline{\mathsf{I}}_h^k u - \underline{\mathsf{u}}_h)$ measured in the $\|\cdot\|_{\sharp,h}$-norm. The proof is postponed to Section 6.2.

Theorem 9 (Error estimate)**.** *Assume $\varsigma \geqslant 1 + \frac{C_{\mathrm{tr}}^2 N_\partial}{2}$, (A1)–(A3), and* (26)*. Denote by u and $\underline{\mathsf{u}}_h$ the unique solutions to* (2) *and* (19)*, respectively, and assume that $u_{|T} \in H^{k+2}(T)$ for all $T \in \mathcal{T}_h$. Then, there exists a real number $\gamma' > 0$ depending on ϱ, d, and k, but independent of h, ν, $\boldsymbol{\beta}$, and μ, such that, letting $\widehat{\underline{\mathsf{u}}}_h := \underline{\mathsf{I}}_h^k u$ and $\zeta = \min_{T \in \mathcal{T}_h}(\frac{1}{2}, \tau_{\mathrm{ref},T}\mu_0)$,*

$$\begin{aligned} &\gamma'\zeta\|\widehat{\underline{\mathsf{u}}}_h - \underline{\mathsf{u}}_h\|_{\sharp,h} \\ &\leqslant \left\{ \sum_{T \in \mathcal{T}_h} (\nu_T \|u\|_{H^{k+2}(T)}^2 + \tau_{\mathrm{ref},T}^{-1} \|u\|_{H^{k+1}(T)}^2) h_T^{2(k+1)} + \beta_{\mathrm{ref},T} \min(1, \mathrm{Pe}_T) h_T^{2k+1} \|u\|_{H^{k+1}(T)}^2 \right\}^{1/2} \end{aligned} \tag{29}$$

where $\mathrm{Pe}_T = \max_{F \in \mathcal{F}_T} \|\mathrm{Pe}_{TF}\|_{L^\infty(F)}$ is a local Péclet number (conventionally, $\|\mathrm{Pe}_{TF}\|_{L^\infty(F)} = +\infty$ if $\nu_F = 0$).

Remark 10 (Regime-dependent estimate)**.** *Using the local Péclet number in* (29) *allows us to establish an error estimate which locally adjusts to the various regimes of* (2)*. In mesh cells where diffusion dominates so that $\mathrm{Pe}_T \leqslant h_T$, the contribution to the right-hand side of* (2) *is $\mathcal{O}(h_T^{2(k+1)})$. In mesh cells where advection dominates so that $\mathrm{Pe}_T \geqslant 1$, the contribution is $\mathcal{O}(h_T^{2k+1})$. The transition region, where Pe_T is between h_T and 1, corresponds to intermediate orders of convergence. Notice also that the diffusive contribution exhibits the superconvergent behavior $\mathcal{O}(h_T^{2(k+1)})$ typical of HHO methods, see [9, 11]. As a result, the balancing with the advective contribution is slightly different with respect to other methods where the diffusive contribution typically scales as $\mathcal{O}(h_T^{2k})$.*

5.4 Link with Hybrid Mixed Mimetic methods

We assume here that the diffusion is not degenerate, i.e. $\underline{\nu} > 0$, and show that, under a slight modification of the definition of Pe_{TF}, see (30) below, the present discontinuous-skeletal method for $k = 0$ corresponds to an edge-based Hybrid Mimetic Mixed (HMM) method studied for advective-diffusive equations in [2]. In this section, we consider that the local Péclet number Pe_{TF} is no longer a function defined on the edge F, but the average of this function, i.e.,

$$\mathrm{Pe}_{TF} = \frac{1}{|F|}\int_F \frac{h_F}{\nu_F}\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF}. \tag{30}$$

With this new definition, and assuming that $\underline{\nu} > 0$, the edge-based HMM method for (2) with $\mu = 0$ can be written (see [2, Eqs. (2.48)–(2.49)]) in the flux balance and continuity form as follows:

$$\forall T \in \mathcal{F}_h \,:\, \sum_{F\in\mathcal{F}_T} |F|\left[(\mathbb{F}_{\mathrm{d}})_{TF} + (\mathbb{F}_{\mathrm{a}})_{TF}\right] = \int_T f \tag{31}$$

$$\forall F \in \mathcal{F}_T \cap \mathcal{F}_{T'} \text{ with } T \neq T' \,:\, (\mathbb{F}_{\mathrm{d}})_{TF} + (\mathbb{F}_{\mathrm{a}})_{TF} + (\mathbb{F}_{\mathrm{d}})_{T'F} + (\mathbb{F}_{\mathrm{a}})_{T'F} = 0, \tag{32}$$

where $\mathbb{F}_{\mathrm{d}}$ and $\mathbb{F}_{\mathrm{a}}$ are diffusion and advection fluxes, constructed from the unknown $\underline{\mathsf{u}}_h \in \underline{\mathsf{U}}_h^0$. We additionally assume that boundary conditions are strongly enforced by considering the space $\underline{\mathsf{U}}_{h,0}^0 := \left\{\underline{\mathsf{v}}_h \in \underline{\mathsf{U}}_h^0 \mid \mathsf{v}_F \equiv 0 \quad \forall F \in \mathcal{F}_h^{\mathrm{b}}\right\}$ (we are entitled to strongly enforce boundary conditions since we assume $\underline{\nu} > 0$ in this section). Taking $\underline{\mathsf{v}}_h \in \underline{\mathsf{U}}_{h,0}^0$, multiplying (31) by the constant value v_T, summing on the cells $T \in \mathcal{T}_h$, and using the flux conservativity (32) and the strong boundary condition to introduce the constant value v_F in the sums, we see that these two equations are equivalent to

$$\sum_{T\in\mathcal{F}_h}\sum_{F\in\mathcal{F}_T} |F|\left[(\mathbb{F}_{\mathrm{d}})_{TF} + (\mathbb{F}_{\mathrm{a}})_{TF}\right](\mathsf{v}_T - \mathsf{v}_F) = l_h(\underline{\mathsf{v}}_h), \tag{33}$$

for all $\underline{\mathsf{v}}_h \in \underline{\mathsf{U}}_{h,0}^0$. As seen in [11], the definition of the diffusive flux $\mathbb{F}_{\mathrm{d}}$ in [15, Eq. (2.25)] shows that, when the stabilization matrices $\mathbb{B}^T$ in the HMM method are diagonal with coefficients $(\frac{\nu_T}{h_F}|F|)_{F\in\mathcal{F}_T}$, the local diffusive term $\sum_{F\in\mathcal{F}_T}|F|(\mathbb{F}_{\mathrm{d}})_{TF}(\mathsf{v}_T - \mathsf{v}_F)$ is identical to the local diffusive bilinear form $a_{\nu,T}$ defined in (10). Therefore, it remains to study the advective term in (33) and see that it corresponds to $a_{\boldsymbol{\beta},0,h}(\underline{\mathsf{u}}_h, \underline{\mathsf{v}}_h)$. With the choice (30), using the diffusive scaling mentioned in [2, §2.4.1] and applying a local geometric scaling based on the edge diameter h_F rather than the distance between the two neighboring cell centers, the advective flux is written [2, Eq. (2.46)]

$$(\mathbb{F}_{\mathrm{a}})_{TF} = \tfrac{\nu_F}{h_F}\left(A^+(\mathrm{Pe}_{TF})\mathsf{u}_T - A^-(\mathrm{Pe}_{TF})\mathsf{u}_F\right).$$

Since $A^+(s) - A^-(s) = s$ and invoking the assumption $\boldsymbol{\nabla}\cdot\boldsymbol{\beta} \equiv 0$, we find that the advective contribution in (33) is

$$\begin{aligned}
&\sum_{T\in\mathcal{F}_h}\sum_{F\in\mathcal{F}_T} |F|\left[\tfrac{\nu_F}{h_F}A^+(\mathrm{Pe}_{TF})\mathsf{u}_T - \tfrac{\nu_F}{h_F}A^-(\mathrm{Pe}_{TF})\mathsf{u}_F\right](\mathsf{v}_T - \mathsf{v}_F)\\
&= \sum_{T\in\mathcal{F}_h}\sum_{F\in\mathcal{F}_T} |F|\left[\tfrac{\nu_F}{h_F}\left(A^+(\mathrm{Pe}_{TF}) - A^-(\mathrm{Pe}_{TF})\right)\mathsf{u}_T + \tfrac{\nu_F}{h_F}A^-(\mathrm{Pe}_{TF})(\mathsf{u}_T - \mathsf{u}_F)\right](\mathsf{v}_T - \mathsf{v}_F)\\
&= \sum_{T\in\mathcal{F}_h}\sum_{F\in\mathcal{F}_T}\left(\int_F \boldsymbol{\beta}\cdot\boldsymbol{n}_{TF}\right)\mathsf{u}_T(\mathsf{v}_T - \mathsf{v}_F) + \sum_{T\in\mathcal{F}_h}\sum_{F\in\mathcal{F}_T}|F|\tfrac{\nu_F}{h_F}A^-(\mathrm{Pe}_{TF})(\mathsf{u}_T - \mathsf{u}_F)(\mathsf{v}_T - \mathsf{v}_F)\\
&= \sum_{T\in\mathcal{F}_h}\mathsf{u}_T\mathsf{v}_T\int_T \boldsymbol{\nabla}\cdot\boldsymbol{\beta} - \sum_{T\in\mathcal{F}_h}\mathsf{u}_T\sum_{F\in\mathcal{F}_T}\left(\int_F \boldsymbol{\beta}\cdot\boldsymbol{n}_{TF}\right)\mathsf{v}_F + s^-_{\boldsymbol{\beta},h}(\underline{\mathsf{u}}_h, \underline{\mathsf{v}}_h)\\
&= -\sum_{T\in\mathcal{F}_h}|T|\mathsf{u}_T\left(\frac{1}{|T|}\sum_{F\in\mathcal{F}_T}\left(\int_F \boldsymbol{\beta}\cdot\boldsymbol{n}_{TF}\right)\mathsf{v}_F\right) + s^-_{\boldsymbol{\beta},h}(\underline{\mathsf{u}}_h, \underline{\mathsf{v}}_h).
\end{aligned}$$

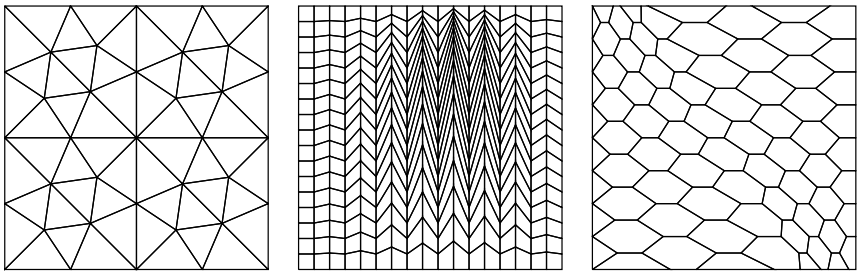


Figure 1: Meshes for the test case of Section 5.5.1. From left to right, the meshes are refereed to as triangular, Kershaw and hexagonal, respectively

It is then a simple matter of inspecting the definition (12a) in the case $k = 0$ to notice that

$$G^k_{\boldsymbol{\beta},T}\underline{\mathsf{v}}_T = \frac{1}{|T|}\sum_{F\in\mathcal{F}_T}\left(\int_F \boldsymbol{\beta}\cdot\boldsymbol{n}_{TF}\right)\mathsf{v}_F,$$

and therefore conclude that the advective contribution in (33) is indeed $a_{\boldsymbol{\beta},0,h}(\underline{\mathsf{u}}_h,\underline{\mathsf{v}}_h)$.

5.5 Numerical results

To close this section, we provide numerical results illustrating the error estimate of Theorem 9.

5.5.1 Uniform diffusion

To numerically assess the sharpness of estimate (29) in the uniform diffusion case, we solve on the unit square the problem (2) with boundary conditions and right-hand side inferred from the following exact solution:

$$u(\boldsymbol{x}) = \sin(\pi x_1)\sin(\pi x_2).$$

We take $\boldsymbol{\beta}(\boldsymbol{x}) = (1/2 - x_2, x_1 - 1/2)$, $\mu = 1$, and we let ν vary in $\{0, 10^{-3}, 1\}$. In Figure 2 we display the convergence results for the three mesh families depicted in Figure 1. From left to right, these correspond, respectively, to the mesh families 1 (triangular) and 4.1 (Kershaw) of the FVCA5 benchmark [20], and to the (predominantly) hexagonal mesh family first introduced in [12]. Each line in Figure 2 corresponds to a different mesh family, and the value of ν increases from left to right. In all the cases, an increase in the asymptotic convergence rate of about half a unit is observed as we increase the value of ν, as predicted by (29). The results also show that the method behaves consistently on a variety of meshes possibly including general polygonal elements. The slightly higher convergence rates for the Kershaw mesh family are possibly due to the fact that the mesh regularity changes when refining.

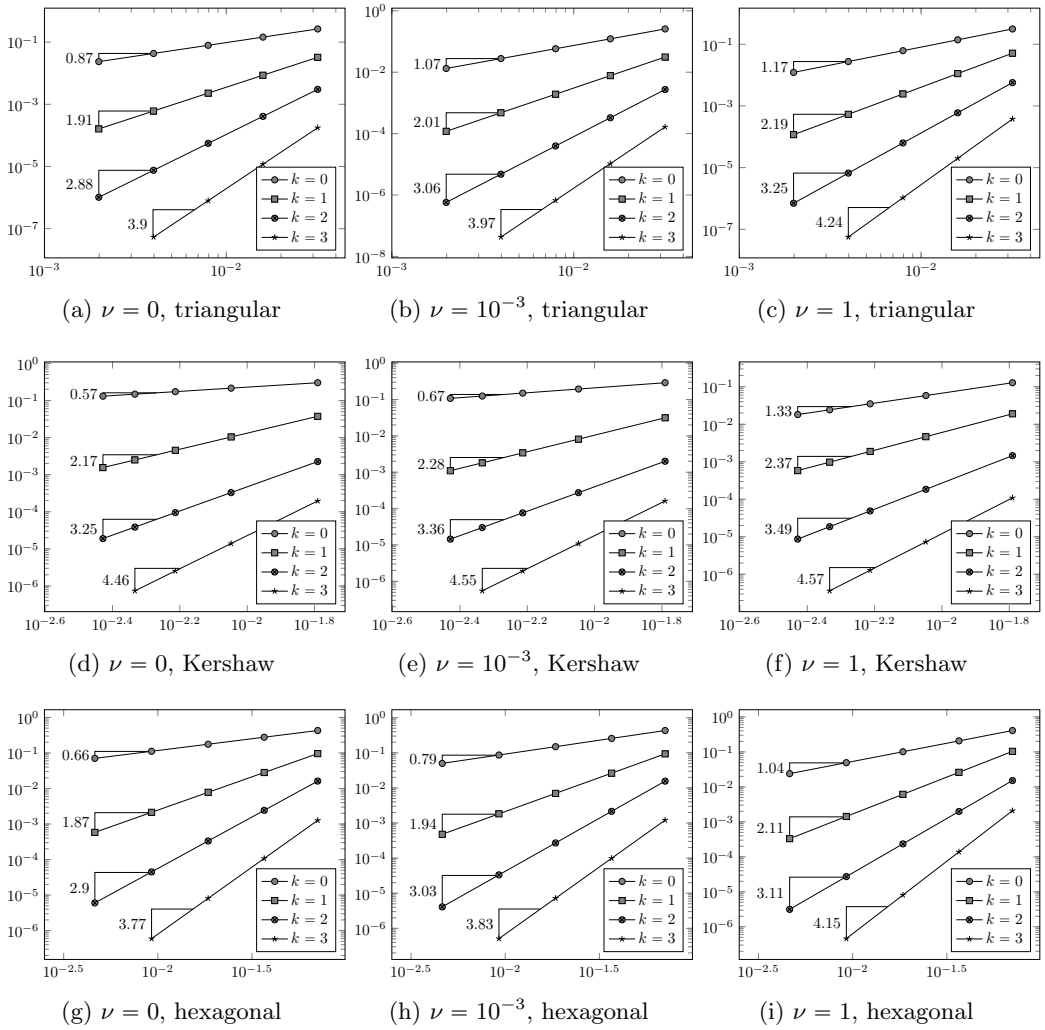
5.5.2 Locally degenerate diffusion

To validate the method in the locally degenerate case, we consider the configuration originally proposed in [10, Section 6.1], cf. Figure 3. The domain is $\Omega = (-1,1)^2\backslash[-0.5,0.5]^2$. Denoting by (r,θ) the standard polar coordinates (with azimuth θ measured anticlockwise starting from the positive x-axis) and by $\boldsymbol{e}_\theta$ the azimuthal vector, the problem coefficients are

$$\nu(\theta,r) = \begin{cases}\pi & \text{if } 0<\theta<\pi,\\ 0 & \text{if } \pi<\theta<2\pi,\end{cases}\qquad \boldsymbol{\beta}(\theta,r) = \frac{\boldsymbol{e}_\theta}{r},\quad \mu = 10^{-6},$$

The exact solution, also used to infer the value of the forcing term f and boundary datum g, is given by

$$u(\theta,r) = \begin{cases}(\theta-\pi)^2 & \text{if } 0<\theta<\pi,\\ 3\pi(\theta-\pi) & \text{if } \pi<\theta<2\pi.\end{cases}$$



(a) $\nu = 0$, triangular (b) $\nu = 10^{-3}$, triangular (c) $\nu = 1$, triangular

(d) $\nu = 0$, Kershaw (e) $\nu = 10^{-3}$, Kershaw (f) $\nu = 1$, Kershaw

(g) $\nu = 0$, hexagonal (h) $\nu = 10^{-3}$, hexagonal (i) $\nu = 1$, hexagonal

Figure 2: $\|\widehat{\underline{\mathrm{u}}}_h - \underline{\mathrm{u}}_h\|_{\sharp,h}$-norm vs. h for different mesh families (rows) and values of the diffusion coefficient ν (columns) in the test case of Section 5.5.1

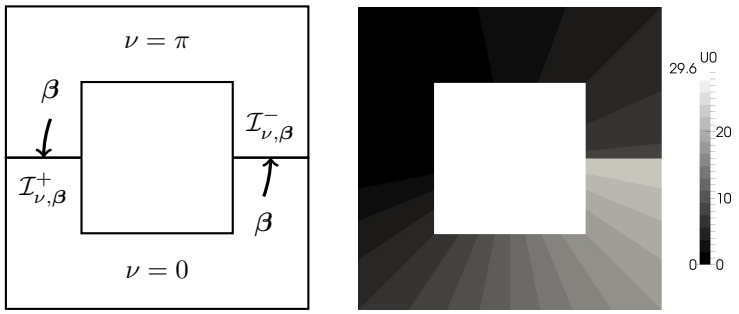
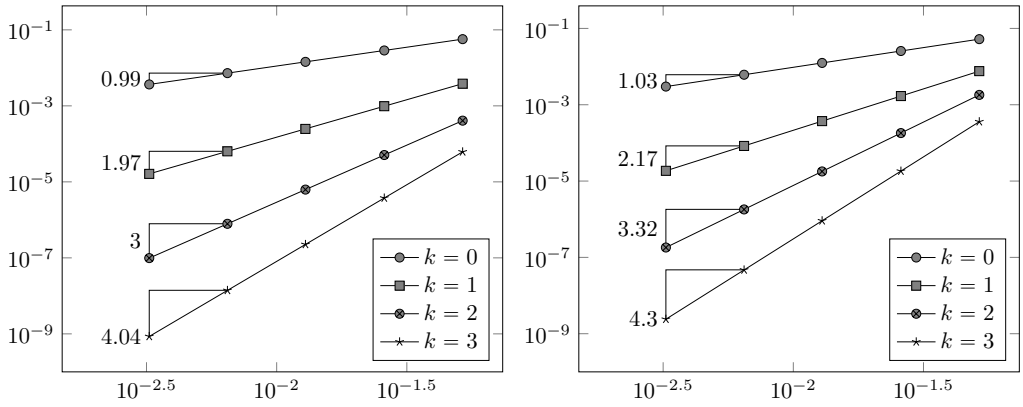


Figure 3: Configuration for the test case of Section 5.5.2 (left) and numerical solution for $k = 3$ and $h = 1.29\times10^{-2}$ (right). The jump discontinuity across $\mathcal{I}^-_{\nu,\boldsymbol{\beta}}$ is clearly visible.



(a) L^2-error on u (b) $\|\widehat{\underline{\mathsf{u}}}_h - \underline{\mathsf{u}}_h\|_{\sharp,h}$

Figure 4: Convergence results for the locally degenerate test case of Section 5.5.2

In Figure 4 we show the convergence results for a refined family of triangular meshes. The left panel displays the L^2-error on the potential measured by the quantity $\{\sum_{T\in\mathcal{T}_h} \|\widehat{\mathsf{u}}_T - \mathsf{u}_T\|_T^2\}^{1/2}$ with $\widehat{\underline{\mathsf{u}}} := \underline{\mathsf{I}}_h^k u$, while the right panel contains the error in the $\|\cdot\|_{\sharp,h}$-norm defined by (24). In both cases the relative error is displayed and we have taken $\varsigma = 1$.

6 Proofs

This section is concerned with the proof of our two main results: Lemma 6 on stability and Theorem 9 on the error estimate. In what follows, we often abbreviate $a \lesssim b$ the inequality $a \leqslant Cb$ with $C > 0$ independent of h, ν, $\boldsymbol{\beta}$, and μ, but possibly depending on ϱ, d, and k.

6.1 Stability analysis

This section is organized as follows. First, we examine separately the coercivity of the diffusive and the advective-reactive bilinear forms. Combining these results readily yields the coercivity of the bilinear form a_h, see (25) in Lemma 6. Then, we prove the inf-sup condition (27).

Lemma 11 (Stability of $a_{\nu,h}$)**.** *Assume $\varsigma \geqslant 1 + \frac{C_{\mathrm{tr}}^2 N_\partial}{2}$. Then, for all $\underline{\mathsf{v}}_h \in \underline{\mathsf{U}}_h^k$, the following holds:*

$$\frac{1}{2}\|\underline{\mathsf{v}}_h\|_{\nu,h}^2 \lesssim a_{\nu,h}(\underline{\mathsf{v}}_h, \underline{\mathsf{v}}_h).$$

Proof. We use the Cauchy–Schwarz and discrete trace (5) inequalities, the definition (20) of the $\|\cdot\|_{\nu,T}$-seminorm, and we recall (4) to obtain

$$\begin{aligned}\left|\sum_{F\in\mathcal{F}_h^{\mathrm{b}}} (\nu_F \boldsymbol{\nabla} p_{T(F)}^k \underline{\mathsf{v}}_{T(F)}\cdot\boldsymbol{n}_{TF}, \mathsf{v}_F)_F\right| &\leqslant \left\{\sum_{F\in\mathcal{F}_h^{\mathrm{b}}} h_F \|\nu_F^{1/2}\boldsymbol{\nabla} p_{T(F)}^k \underline{\mathsf{v}}_{T(F)}\|_F^2\right\}^{1/2} |\underline{\mathsf{v}}_h|_{\nu,\partial\Omega} \\ &\leqslant C_{\mathrm{tr}} N_\partial^{1/2} \left\{\sum_{T\in\mathcal{T}_h} \|\underline{\mathsf{v}}_T\|_{\nu,T}^2\right\}^{1/2} |\underline{\mathsf{v}}_h|_{\nu,\partial\Omega}.\end{aligned}$$

Hence,

$$a_h(\underline{\mathsf{v}}_h, \underline{\mathsf{v}}_h) \geqslant \sum_{T\in\mathcal{T}_h} \|\underline{\mathsf{v}}_T\|_{\nu,T}^2 - C_{\mathrm{tr}} N_\partial^{1/2} \left\{\sum_{T\in\mathcal{T}_h} \|\underline{\mathsf{v}}_T\|_{\nu,T}^2\right\}^{1/2} |\underline{\mathsf{v}}_h|_{\nu,\partial\Omega} + \varsigma |\underline{\mathsf{v}}_h|_{\nu,\partial\Omega}^2. \tag{34}$$

For a real number $A > 0$, assuming $B > A^2$, the following inequality holds for all positive real numbers x, y: $x^2 - 2Axy + By^2 \geqslant \frac{B-A^2}{1+B}(x^2+y^2)$. Using this result with $x = \left\{\sum_{T\in\mathcal{T}_h} \|\underline{\mathsf{v}}_T\|_{\nu,T}^2\right\}^{1/2}$, $y = |\underline{\mathsf{v}}_h|_{\nu,\partial\Omega}$, $2A = C_{\rm tr}N_\partial^{1/2}$, and $B = \varsigma$ in the right-hand side of (34) yields the assertion since $\frac{B-A^2}{1+B} \geqslant \frac{1}{2}$. □

Lemma 12 (Stability of $a_{\boldsymbol{\beta},\mu,h}$)**.** *Assume (A1)–(A3). The following holds for all $\underline{\mathsf{v}}_h \in \underline{\mathsf{U}}_h^k$:*

$$\eta\|\underline{\mathsf{v}}_h\|_{\boldsymbol{\beta},\mu,h}^2 \leqslant a_{\boldsymbol{\beta},\mu,h}(\underline{\mathsf{v}}_h,\underline{\mathsf{v}}_h),$$

with $\eta := \min_{T\in\mathcal{T}_h}(1,\tau_{{\rm ref},T}\mu_0)$.

Proof. Step 1. Let us first prove that, for all $\underline{\mathsf{w}}_h, \underline{\mathsf{v}}_h \in \underline{\mathsf{U}}_h^k$, the following holds:

$$\sum_{T\in\mathcal{T}_h}\Big\{(G_{\boldsymbol{\beta},T}^k\underline{\mathsf{w}}_T,\mathsf{v}_T)_T + (\mathsf{w}_T,G_{\boldsymbol{\beta},T}^k\underline{\mathsf{v}}_T)_T\Big\} = - \sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_h}((\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF})(\mathsf{w}_F-\mathsf{w}_T),\mathsf{v}_F-\mathsf{v}_T)_F + \sum_{F\in\mathcal{F}_h^{\rm b}}((\boldsymbol{\beta}\cdot\boldsymbol{n}_F)\mathsf{w}_F,\mathsf{v}_F)_F. \quad (35)$$

For all $T \in \mathcal{T}_h$, we use (12b) with $\underline{\mathsf{v}}_T = \underline{\mathsf{w}}_T$ and $w = \mathsf{v}_T$ and (12a) with $w = \mathsf{w}_T$ to infer that

$$\begin{aligned}
&\sum_{T\in\mathcal{T}_h}(G_{\boldsymbol{\beta},T}^k\underline{\mathsf{w}}_T,\mathsf{v}_T)_T\\
&\quad= \sum_{T\in\mathcal{T}_h}\left\{(\boldsymbol{\beta}\cdot\boldsymbol{\nabla}\mathsf{w}_T,\mathsf{v}_T)_T + \sum_{F\in\mathcal{F}_T}((\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF})(\mathsf{w}_F-\mathsf{w}_T),\mathsf{v}_T)_F\right\}\\
&\quad= \sum_{T\in\mathcal{T}_h}\left\{-(\mathsf{w}_T,G_{\boldsymbol{\beta},T}^k\underline{\mathsf{v}}_T)_T + \sum_{F\in\mathcal{F}_T}((\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF})\mathsf{w}_T,\mathsf{v}_F)_F + \sum_{F\in\mathcal{F}_T}((\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF})(\mathsf{w}_F-\mathsf{w}_T),\mathsf{v}_T)_F\right\}.
\end{aligned} \quad (36)$$

Formula (35) follows adding to the right-hand side of (36) the quantity

$$0 = \sum_{F\in\mathcal{F}_h^{\rm b}}((\boldsymbol{\beta}\cdot\boldsymbol{n}_F)\mathsf{w}_F,\mathsf{v}_F)_F - \sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T}((\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF})\mathsf{w}_F,\mathsf{v}_F)_F. \quad (37)$$

To prove (37), we observe that, rearranging the sums,

$$\sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T}((\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF})\mathsf{w}_F,\mathsf{v}_F)_F = \sum_{F\in\mathcal{F}_h^{\rm i}}\sum_{T\in\mathcal{T}_F}((\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF})\mathsf{w}_F,\mathsf{v}_F)_F + \sum_{F\in\mathcal{F}_h^{\rm b}}((\boldsymbol{\beta}\cdot\boldsymbol{n}_F)\mathsf{w}_F,\mathsf{v}_F)_F.$$

Using, for all $F \in \mathcal{F}_h^{\rm i}$ such that $F \subset \partial T_1 \cap \partial T_2$, the fact that $\boldsymbol{\beta}\cdot\boldsymbol{n}_{T_1F} = -\boldsymbol{\beta}\cdot\boldsymbol{n}_{T_2F} = \boldsymbol{\beta}\cdot\boldsymbol{n}_F$, we infer that the first addend in the right-hand side is zero.

Step 2. Owing to (13) (see also (14) if $\nu_F = 0$) and since $A^+(s) - A^-(s) = s$, we observe that, for all $T \in \mathcal{T}_h$ and all $F \in \mathcal{F}_T$,

$$\tfrac{\nu_F}{h_F}A^+({\rm Pe}_{TF}) - \tfrac{\nu_F}{h_F}A^-({\rm Pe}_{TF}) = \boldsymbol{\beta}\cdot\boldsymbol{n}_{TF}. \quad (38)$$

Owing to (35), we infer that, for all $\underline{\mathsf{w}}_h, \underline{\mathsf{v}}_h \in \underline{\mathsf{U}}_h^k$,

$$\begin{aligned}
&a_{\boldsymbol{\beta},\mu,h}(\underline{\mathsf{w}}_h,\underline{\mathsf{v}}_h)\\
&\quad= \sum_{T\in\mathcal{T}_h}\left\{-(\mathsf{w}_T,G_{\boldsymbol{\beta},T}^k\underline{\mathsf{v}}_T)_T + (\mu\mathsf{w}_T,\mathsf{v}_T)_T\right\} + s_{\boldsymbol{\beta},h}^-(\underline{\mathsf{w}}_h,\underline{\mathsf{v}}_h) + \sum_{F\in\mathcal{F}_h^{\rm b}}(\tfrac{\nu_F}{h_F}A^+({\rm Pe}_F)\mathsf{w}_F,\mathsf{v}_F)_F \quad (39a)\\
&\quad= \sum_{T\in\mathcal{T}_h}\left\{(G_{\boldsymbol{\beta},T}^k\underline{\mathsf{w}}_T,\mathsf{v}_T)_T + (\mu\mathsf{w}_T,\mathsf{v}_T)_T\right\} + s_{\boldsymbol{\beta},h}^+(\underline{\mathsf{w}}_h,\underline{\mathsf{v}}_h) + \sum_{F\in\mathcal{F}_h^{\rm b}}(\tfrac{\nu_F}{h_F}A^-({\rm Pe}_F)\mathsf{w}_F,\mathsf{v}_F)_F, \quad (39b)
\end{aligned}$$

where the global stabilization bilinear forms $s^{\pm}_{\boldsymbol{\beta},h}$ on $\underline{\mathsf{U}}^k_h \times \underline{\mathsf{U}}^k_h$ are assembled element-wise by setting $s^{\pm}_{\boldsymbol{\beta},h}(\underline{\mathsf{w}}_h, \underline{\mathsf{v}}_h) := \sum_{T\in\mathcal{T}_h} s^{\pm}_{\boldsymbol{\beta},T}(\underline{\mathsf{w}}_T, \underline{\mathsf{v}}_T)$.
Step 3. Let $\underline{\mathsf{v}}_h \in \underline{\mathsf{U}}^k_h$. Using (35) with $\underline{\mathsf{w}}_h = \underline{\mathsf{v}}_h$ and (38), we infer that

$$\begin{aligned} -\sum_{T\in\mathcal{T}_h} (G^k_{\boldsymbol{\beta},T}\underline{\mathsf{v}}_T, \mathsf{v}_T)_T = \sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} \left(\frac{\frac{\nu_F}{h_F}A^+(\mathrm{Pe}_{TF}) - \frac{\nu_F}{h_F}A^-(\mathrm{Pe}_{TF})}{2}(\mathsf{v}_F - \mathsf{v}_T), \mathsf{v}_F - \mathsf{v}_T \right)_F \\ - \sum_{F\in\mathcal{F}^{\mathrm{b}}_h} \left(\frac{\frac{\nu_F}{h_F}A^+(\mathrm{Pe}_F) - \frac{\nu_F}{h_F}A^-(\mathrm{Pe}_F)}{2}\mathsf{v}_F, \mathsf{v}_F \right)_F. \end{aligned} \tag{40}$$

Taking $\underline{\mathsf{w}}_h = \underline{\mathsf{v}}_h$ in (39a) and using (40) to substitute the first term in the right-hand side, we obtain

$$\begin{aligned} a_{\boldsymbol{\beta},\mu,h}(\underline{\mathsf{v}}_h, \underline{\mathsf{v}}_h) \geqslant \sum_{T\in\mathcal{T}_h} \left\{ \sum_{F\in\mathcal{F}_T} \left(\frac{\frac{\nu_F}{h_F}A^+(\mathrm{Pe}_{TF}) + \frac{\nu_F}{h_F}A^-(\mathrm{Pe}_{TF})}{2}(\mathsf{v}_F - \mathsf{v}_T), \mathsf{v}_F - \mathsf{v}_T \right)_F + \mu_0 \|\mathsf{v}_T\|^2_T \right\} \\ + \frac{1}{2}\sum_{F\in\mathcal{F}^{\mathrm{b}}_h} \left(\frac{\frac{\nu_F}{h_F}A^+(\mathrm{Pe}_F) + \frac{\nu_F}{h_F}A^-(\mathrm{Pe}_F)}{2}\mathsf{v}_F, \mathsf{v}_F \right)_F, \end{aligned}$$

and the conclusion follows recalling (23) and since $|A|(s) = A^+(s) + A^-(s)$. □

Proof of (25). Sum the results of Lemmas 11 and 12. □

Proof of the inf-sup condition (27). The proof hinges on the use of the locally scaled advective derivative as a test function, an idea which can be found, e.g., in the work of Johnson and Pitkäranta [22]. We denote by \$ the supremum in the right-hand side of (27). Let $\underline{\mathsf{w}}_h \in \underline{\mathsf{U}}^k_h$ and define $\underline{\mathsf{v}}_h \in \underline{\mathsf{U}}^k_h$ such that,

$$\mathsf{v}_T = h_T\beta^{-1}_{\mathrm{ref},T}(G^k_{\boldsymbol{\beta},T}\underline{\mathsf{w}}_T) \quad \forall T\in\mathcal{T}_h, \qquad \mathsf{v}_F \equiv 0 \quad \forall F\in\mathcal{F}_h. \tag{41}$$

The following result is proved in Lemma 13:

$$\|\underline{\mathsf{v}}_h\|_{\sharp,h} \lesssim \|\underline{\mathsf{w}}_h\|_{\sharp,h}. \tag{42}$$

Using (41) in (17) and recalling (39b), it is inferred that

$$\begin{aligned} \sum_{T\in\mathcal{T}_h} h_T\beta^{-1}_{\mathrm{ref},T}\|G^k_{\boldsymbol{\beta},T}\underline{\mathsf{w}}_T\|^2_T = a_h(\underline{\mathsf{w}}_h, \underline{\mathsf{v}}_h) - a_{\nu,h}(\underline{\mathsf{w}}_h, \underline{\mathsf{v}}_h) - \sum_{T\in\mathcal{T}_h}(\mu\mathsf{w}_T, \mathsf{v}_T)_T \\ - s^+_{\boldsymbol{\beta},h}(\underline{\mathsf{w}}_h, \underline{\mathsf{v}}_h) - \sum_{F\in\mathcal{F}^{\mathrm{b}}_h}(\tfrac{\nu_F}{h_F}A^-(\mathrm{Pe}_F)\mathsf{w}_F, \mathsf{v}_F)_F. \end{aligned} \tag{43}$$

Denote by $\mathfrak{T}_1, \ldots, \mathfrak{T}_5$ the addends in the right-hand side of (43). Using (42), we have

$$|\mathfrak{T}_1| \leqslant \$\|\underline{\mathsf{v}}_h\|_{\sharp,h} \lesssim \$\|\underline{\mathsf{w}}_h\|_{\sharp,h}. \tag{44}$$

Since $\mathsf{v}_F = 0$ for any face F, using the Cauchy-Schwarz inequality on the positive semi-definite bilinear form $a_{\nu,T}$ and recalling the definition (20) of $\|\cdot\|_{\nu,h}$, it is inferred, thanks to (42), that

$$|\mathfrak{T}_2| \leqslant \|\underline{\mathsf{w}}_h\|_{\nu,h}\|\underline{\mathsf{v}}_h\|_{\nu,h} \lesssim \|\underline{\mathsf{w}}_h\|_{\flat,h}\|\underline{\mathsf{w}}_h\|_{\sharp,h}. \tag{45}$$

The estimate on $\mathfrak{T}_3$ is trivial:

$$|\mathfrak{T}_3| \lesssim \|\underline{\mathsf{w}}_h\|_{\boldsymbol{\beta},\mu,h}\|\underline{\mathsf{v}}_h\|_{\boldsymbol{\beta},\mu,h} \lesssim \|\underline{\mathsf{w}}_h\|_{\flat,h}\|\underline{\mathsf{w}}_h\|_{\sharp,h}. \tag{46}$$

Let us now turn to $\mathfrak{T}_4$. Using Remark 2 (if $\nu_F > 0$) and (A3) (otherwise) we see that

$$|\tfrac{\nu_F}{h_F}A^+(\mathrm{Pe}_{TF})| \lesssim \frac{\nu_F}{h_F} + \frac{\nu_F}{h_F}|A|(\mathrm{Pe}_{TF}).$$

Using the fact that $\nu_F \leqslant \nu_T$ and $h_T \lesssim h_F$ owing to (3) whenever $F \in \mathcal{F}_T$, the norm equivalence (21), the Cauchy–Schwarz inequality, and the definition (22) of the advective seminorm $\|\cdot\|_{\boldsymbol{\beta},\mu,h}$, we therefore find

$$\begin{aligned}
|\mathfrak{T}_4| &\leqslant \sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} (|\tfrac{\nu_F}{h_F}A^+(\mathrm{Pe}_{TF})|\,|\mathsf{w}_F-\mathsf{w}_T|,|\mathsf{v}_F-\mathsf{v}_T|)_F\\
&\lesssim \sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} \tfrac{\nu_T}{h_T}(|\mathsf{w}_F-\mathsf{w}_T|,|\mathsf{v}_F-\mathsf{v}_T|)_F\\
&\quad + \sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} (\tfrac{\nu_F}{h_F}|A|(\mathrm{Pe}_{TF})\,|\mathsf{w}_F-\mathsf{w}_T|,|\mathsf{v}_F-\mathsf{v}_T|)_F\\
&\lesssim \|\underline{\mathsf{w}}_h\|_{\nu,h}\|\underline{\mathsf{v}}_h\|_{\nu,h} + \|\underline{\mathsf{w}}_h\|_{\boldsymbol{\beta},\mu,h}\|\underline{\mathsf{v}}_h\|_{\boldsymbol{\beta},\mu,h} \lesssim \|\underline{\mathsf{w}}_h\|_{\flat,h}\|\underline{\mathsf{v}}_h\|_{\sharp,h}.
\end{aligned} \tag{47}$$

Proceeding similarly, it is inferred for $\mathfrak{T}_5$:

$$|\mathfrak{T}_5| \lesssim \|\underline{\mathsf{w}}_h\|_{\flat,h}\|\underline{\mathsf{w}}_h\|_{\sharp,h}. \tag{48}$$

Hence, using (44)–(48) to bound the right-hand side of (43), we obtain

$$\sum_{T\in\mathcal{T}_h} h_T\beta_{\mathrm{ref},T}^{-1}\|G^k_{\boldsymbol{\beta},T}\underline{\mathsf{w}}_T\|_T^2 \lesssim \$\|\underline{\mathsf{w}}_h\|_{\sharp,h} + \|\underline{\mathsf{w}}_h\|_{\flat,h}\|\underline{\mathsf{w}}_h\|_{\sharp,h}. \tag{49}$$

Adding $\|\underline{\mathsf{w}}_h\|_{\flat,h}^2$ to both sides of inequality (49), and observing that, as a consequence of (25),

$$\|\underline{\mathsf{w}}_h\|_{\flat,h}^2 \leqslant \zeta^{-1}\frac{a_h(\underline{\mathsf{w}}_h,\underline{\mathsf{w}}_h)}{\|\underline{\mathsf{w}}_h\|_{\sharp,h}}\|\underline{\mathsf{w}}_h\|_{\sharp,h} \leqslant \zeta^{-1}\$\|\underline{\mathsf{w}}_h\|_{\sharp,h}, \tag{50}$$

we infer the existence of C depending on ϱ, d, and k, but independent of h, ν, $\boldsymbol{\beta}$, and μ, such that

$$C\|\underline{\mathsf{w}}_h\|_{\sharp,h}^2 \leqslant \zeta^{-1}\$\|\underline{\mathsf{w}}_h\|_{\sharp,h} + \|\underline{\mathsf{w}}_h\|_{\flat,h}\|\underline{\mathsf{w}}_h\|_{\sharp,h} \leqslant \zeta^{-1}\$\|\underline{\mathsf{w}}_h\|_{\sharp,h} + \frac{1}{2C}\|\underline{\mathsf{w}}_h\|_{\flat,h}^2 + \frac{C}{2}\|\underline{\mathsf{w}}_h\|_{\sharp,h}^2,$$

and the result follows using again (50) for the second term in the right-hand side. □

Lemma 13. *Under the assumptions of Lemma 6, let $\underline{\mathsf{w}}_h \in \underline{\mathsf{U}}_h^k$ and $\underline{\mathsf{v}}_h \in \underline{\mathsf{U}}_h^k$ be defined as in* (41)*. Then,* (42) *holds.*

Proof. Using (12b), we observe that, for all $\underline{\mathsf{z}}_T \in \underline{\mathsf{U}}_T^k$,

$$\begin{aligned}
\sqrt{\nu_T}\|G^k_{\boldsymbol{\beta},T}\underline{\mathsf{z}}_T\|_T &= \sup_{w\in\mathbb{P}_d^k(T),\|w\|_T=1}\sqrt{\nu_T}(G^k_{\boldsymbol{\beta},T}\underline{\mathsf{z}}_T,w)_T\\
&\lesssim \left\{\nu_T\|\boldsymbol{\beta}\cdot\boldsymbol{\nabla}\mathsf{z}_T\|_T^2 + \sum_{F\in\mathcal{F}_T}\tfrac{\nu_T}{h_F}\||\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF}|(\mathsf{z}_F-\mathsf{z}_T)\|^2\right\}^{1/2} \lesssim \beta_{\mathrm{ref},T}\|\underline{\mathsf{z}}_T\|_{\nu,T}.
\end{aligned} \tag{51}$$

The first inequality results from multiple applications of the Cauchy–Schwarz inequality together with the discrete trace inequality (5) and the bound (4) on N_∂, while the second is an immediate consequence of the definition (23) of $\beta_{\mathrm{ref},T}$ and of the equivalence (21).

(i) *Diffusive contribution.* Recalling (21), using the discrete inverse (6) and trace (5) inequalities followed by (3) to write $h_T/h_F \leqslant \varrho^{-2}$ and the bound (4) on N_∂ for the boundary term, it is inferred that

$$\begin{aligned}
\|\underline{\mathsf{v}}_h\|_{\nu,h}^2 &\lesssim \sum_{T\in\mathcal{T}_h}\left\{\nu_T h_T^2\beta_{\mathrm{ref},T}^{-2}\|\boldsymbol{\nabla}G^k_{\boldsymbol{\beta},T}\underline{\mathsf{w}}_T\|_T^2 + \sum_{F\in\mathcal{F}_T}\nu_T h_T\beta_{\mathrm{ref},T}^{-2}\|G^k_{\boldsymbol{\beta},T}\underline{\mathsf{w}}_T\|_F^2\right\}\\
&\lesssim \sum_{T\in\mathcal{T}_h}\nu_T\beta_{\mathrm{ref},T}^{-2}\|G^k_{\boldsymbol{\beta},T}\underline{\mathsf{w}}_T\|_T^2 \lesssim \|\underline{\mathsf{w}}_h\|_{\nu,h}^2,
\end{aligned} \tag{52}$$

where, for all $T \in \mathcal{T}_h$, we have used (51) with $\underline{\mathsf{z}}_T = \underline{\mathsf{w}}_T$ to conclude.

(ii) *Advective and reactive contributions.* If $\nu_F > 0$ then, since $|A|$ is Lipschitz-continuous and vanishes at 0,

$$\tfrac{\nu_F}{h_F}|A|(\mathrm{Pe}_{TF}) \lesssim \tfrac{\nu_F}{h_F}|\mathrm{Pe}_{TF}| = |\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF}| \leqslant \beta_{\mathrm{ref},T}.$$

Owing to (A3), this inequality is also valid in the case $\nu_F = 0$. Hence, recalling the definition (41) of $\underline{\mathsf{v}}_h$ and using the discrete trace inequality (5), it is inferred, for all $T \in \mathcal{T}_h$ and all $F \in \mathcal{F}_T$,

$$\|\big[\tfrac{\nu_F}{h_F}|A|(\mathrm{Pe}_{TF})\big]^{1/2}(\mathsf{v}_F - \mathsf{v}_T)\|_F = \|\big[\tfrac{\nu_F}{h_F}|A|(\mathrm{Pe}_{TF})\big]^{1/2}\mathsf{v}_T\|_F \lesssim h_T^{1/2}\beta_{\mathrm{ref},T}^{-1/2}\|G^k_{\boldsymbol{\beta},T}\underline{\mathsf{w}}_T\|_T. \tag{53}$$

Using (53) together with the uniform bound (4) on N_∂ and the definition (41) of $\underline{\mathsf{v}}_h$, we deduce that

$$\begin{aligned}\|\underline{\mathsf{v}}_h\|^2_{\boldsymbol{\beta},\mu,h} &\lesssim \sum_{T\in\mathcal{T}_h}\Big\{h_T\beta_{\mathrm{ref},T}^{-1}\|G^k_{\boldsymbol{\beta},T}\underline{\mathsf{w}}_T\|_T^2 + h_T^2\tau_{\mathrm{ref},T}^{-1}\beta_{\mathrm{ref},T}^{-2}\|G^k_{\boldsymbol{\beta},T}\underline{\mathsf{w}}_T\|_T^2\Big\}\\ &\lesssim \sum_{T\in\mathcal{T}_h} h_T\beta_{\mathrm{ref},T}^{-1}\|G^k_{\boldsymbol{\beta},T}\underline{\mathsf{w}}_T\|_T^2,\end{aligned} \tag{54}$$

where the conclusion follows by noticing that (26) yields $h_T\beta_{\mathrm{ref},T}^{-1}\tau_{\mathrm{ref},T}^{-1} \leqslant 1$. Moreover, recalling (12a) and using the Cauchy–Schwarz and inverse (6) inequalities together with the definition (23) of $\beta_{\mathrm{ref},T}$ to infer $|(\mathsf{v}_T, \boldsymbol{\beta}\cdot\boldsymbol{\nabla}w)_T| \leqslant \|\mathsf{v}_T\|_T\beta_{\mathrm{ref},T}C_{\mathrm{inv}}h_T^{-1}\|w\|_T$, one has, for all $T \in \mathcal{T}_h$,

$$\|G^k_{\boldsymbol{\beta},T}\underline{\mathsf{v}}_T\|_T = \sup_{w\in\mathbb{P}^k_d(T),\|w\|_T=1} -(\mathsf{v}_T, \boldsymbol{\beta}\cdot\boldsymbol{\nabla}w)_T \lesssim \beta_{\mathrm{ref},T}h_T^{-1}\|\mathsf{v}_T\|_T = \|G^k_{\boldsymbol{\beta},T}\underline{\mathsf{w}}_T\|_T, \tag{55}$$

where we have used the definition (41) of v_T to conclude. Hence, using (54) and (55), we estimate the advective and reactive contributions to $\|\underline{\mathsf{v}}_h\|_{\sharp,h}$ as follows:

$$\|\underline{\mathsf{v}}_h\|^2_{\boldsymbol{\beta},\mu,h} + \sum_{T\in\mathcal{T}_h} h_T\beta_{\mathrm{ref},T}^{-1}\|G^k_{\boldsymbol{\beta},T}\underline{\mathsf{v}}_T\|_T^2 \lesssim \sum_{T\in\mathcal{T}_h} h_T\beta_{\mathrm{ref},T}^{-1}\|G^k_{\boldsymbol{\beta},T}\underline{\mathsf{w}}_T\|_T^2. \tag{56}$$

The conclusion then follows from (24) recalling (52) and (56).

□

6.2 Error analysis

We prove here Theorem 9. Owing to (27), we infer that

$$\|\widehat{\underline{\mathsf{u}}}_h - \underline{\mathsf{u}}_h\|_{\sharp,h} \leqslant (\gamma\zeta)^{-1}\sup_{\underline{\mathsf{v}}_h\in\underline{\mathsf{U}}_h^k\setminus\{0\}}\frac{\mathcal{E}_h(\underline{\mathsf{v}}_h)}{\|\underline{\mathsf{v}}_h\|_{\sharp,h}}, \tag{57}$$

where

$$\mathcal{E}_h(\underline{\mathsf{v}}_h) := a_h(\widehat{\underline{\mathsf{u}}}_h - \underline{\mathsf{u}}_h, \underline{\mathsf{v}}_h) = a_h(\widehat{\underline{\mathsf{u}}}_h, \underline{\mathsf{v}}_h) - l_h(\underline{\mathsf{v}}_h) = a_{\nu,h}(\widehat{\underline{\mathsf{u}}}_h, \underline{\mathsf{v}}_h) + a_{\boldsymbol{\beta},\mu,h}(\widehat{\underline{\mathsf{u}}}_h, \underline{\mathsf{v}}_h) - l_h(\underline{\mathsf{v}}_h),$$

is the consistency error. We derive a bound for this quantity for a generic $\underline{\mathsf{v}}_h \in \underline{\mathsf{U}}_h^k$ proceeding in the same spirit as [11, Theorem 8]. Recalling that $f = \boldsymbol{\nabla}\cdot(-\nu\boldsymbol{\nabla}u + \boldsymbol{\beta}u) + \mu u$ a.e. in Ω, we perform an element-by-element integration by parts on the first term in the definition (18) of $l_h(\underline{\mathsf{v}}_h)$. We then use the conservation property

$$(-\nu\boldsymbol{\nabla}u + \boldsymbol{\beta}u)_{|T_1}\cdot\boldsymbol{n}_{T_1F} + (-\nu\boldsymbol{\nabla}u + \boldsymbol{\beta}u)_{|T_2}\cdot\boldsymbol{n}_{T_2F} = 0$$

valid for any interface $F \subset \partial T_1 \cap \partial T_2$ to introduce v_F in the resulting sums. We also notice that, for any face $F \in \mathcal{F}_h^{\mathrm{b}}$, $\frac{\nu_F}{h_F}A^-(\mathrm{Pe}_F)g = \frac{\nu_F}{h_F}A^-(\mathrm{Pe}_F)u$ on F, which results from the boundary condition (2d) if $\nu_F > 0$ and from the definition (14) if $\nu_F = 0$. Letting $\breve{u}_T := p_T^k\widehat{\underline{\mathsf{u}}}_T$ and using the

definitions (10) and (15) for $a_{\nu,h}$, and (39a) and (12b) for $a_{\boldsymbol{\beta},\mu,h}$, we then find

$$\begin{aligned}
&\mathcal{E}_h(\underline{\mathsf{v}}_h) = \\
&\sum_{T\in\mathcal{T}_h}\left\{(\nu_T\boldsymbol{\nabla}(\breve{u}_T-u,\boldsymbol{\nabla}\mathsf{v}_T)_T+\sum_{F\in\mathcal{F}_T}(\nu_T\boldsymbol{\nabla}(\breve{u}_T-u_{|T})\cdot\boldsymbol{n}_{TF},\mathsf{v}_F-\mathsf{v}_T)_F+s_{\nu,T}(\widehat{\underline{\mathsf{u}}}_h,\underline{\mathsf{v}}_h)\right\}\\
&+\sum_{T\in\mathcal{T}_h}\left\{(u-\widehat{\mathsf{u}}_T,\boldsymbol{\beta}\cdot\boldsymbol{\nabla}\mathsf{v}_T+\mu\mathsf{v}_T)_T+\sum_{F\in\mathcal{F}_T}((\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF})(u_{|T}-\widehat{\mathsf{u}}_T),\mathsf{v}_F-\mathsf{v}_T)_F+s^-_{\boldsymbol{\beta},T}(\widehat{\underline{\mathsf{u}}}_h,\underline{\mathsf{v}}_h)\right\}\\
&+\sum_{F\in\mathcal{F}_h^{\mathrm{b}}}\left\{(\nu_F(\boldsymbol{\nabla}(u-\breve{u}_T)\cdot\boldsymbol{n}_{TF},\mathsf{v}_F)_F+\tfrac{\nu_F}{h_F}(A^+(\mathrm{Pe}_F)(\widehat{\mathsf{u}}_F-u),\mathsf{v}_F)_F\right\}.
\end{aligned}\tag{58}$$

We have used the fact that $\sum_{F\in\mathcal{F}_h^{\mathrm{b}}}\frac{c\nu_F}{h_F}(\widehat{\mathsf{u}}_F-g,\mathsf{v}_F)_F=0$. Indeed, for all $F\in\mathcal{F}_h^{\mathrm{b}}$, either $\nu_F=0$ and the corresponding addend vanishes, or $\nu_F>0$ so that $F\subset\Gamma_{\nu,\boldsymbol{\beta}}$ (cf. (1a)) and hence $\widehat{\mathsf{u}}_F=\pi_F^k g$ owing to (2d) and $(\widehat{\mathsf{u}}_F-g,\mathsf{v}_F)_F=(\pi_F^k g-g,\mathsf{v}_F)_F=0$ since $\mathsf{v}_F\in\mathbb{P}_{d-1}^k(F)$.

Denote by $\mathfrak{T}_1$, $\mathfrak{T}_2$, $\mathfrak{T}_3$ the lines composing the right-hand side of (58) and corresponding, respectively, to diffusive terms, advective terms, and weakly enforced boundary conditions.

(i) *Diffusive terms.* Proceeding as in the proof of [11, Theorem 8] yields

$$|\mathfrak{T}_1|\lesssim\left\{\sum_{T\in\mathcal{T}_h}\nu_T h_T^{2(k+1)}\|u\|^2_{H^{k+2}(T)}\right\}^{1/2}\|\underline{\mathsf{v}}_h\|_{\nu,h}.\tag{59}$$

Observe that, to obtain (59), a crucial point is the choice to interpolate $\widehat{\mathsf{u}}_F$ from the diffusive side whenever $F\subset\mathcal{I}^-_{\nu,\boldsymbol{\beta}}$ since this guarantees that $\breve{u}_T$ enjoys the approximation properties (9) whenever $\nu_T\neq0$.

(ii) *Advective-reactive terms.* Denote by $\mathfrak{T}_{2,1}$, $\mathfrak{T}_{2,2}$, and $\mathfrak{T}_{2,3}$ the three addends that compose $\mathfrak{T}_2$. For the first term, observing that $(\pi_T^0\boldsymbol{\beta})\cdot\boldsymbol{\nabla}\mathsf{v}_T\in\mathbb{P}_d^{k-1}(T)\subset\mathbb{P}_d^k(T)$ and recalling that, owing to (28), $\widehat{\mathsf{u}}_T=\pi_T^k u$, we infer that $\mathfrak{T}_{2,1}=\sum_{T\in\mathcal{T}_h}(u-\pi_T^k u,(\boldsymbol{\beta}-\pi_T^0\boldsymbol{\beta})\cdot\boldsymbol{\nabla}\mathsf{v}_T)_T$. Hence,

$$\begin{aligned}
|\mathfrak{T}_{2,1}|&\lesssim\sum_{T\in\mathcal{T}_h}\left\{\|\boldsymbol{\beta}-\pi_T^0\boldsymbol{\beta}\|_{L^\infty(T)^d}\|u-\pi_T^k u\|_T\|\boldsymbol{\nabla}\mathsf{v}_T\|_T+\|\mu\|_{L^\infty(T)}\|u-\pi_T^k u\|_T\|\mathsf{v}_T\|_T\right\}\\
&\lesssim\left\{\sum_{T\in\mathcal{T}_h}\tau_{\mathrm{ref},T}^{-1}h_T^{2(k+1)}\|u\|^2_{H^{k+1}(T)}\right\}^{1/2}\|\underline{\mathsf{v}}_h\|_{\boldsymbol{\beta},\mu,h},
\end{aligned}\tag{60}$$

where the second inequality is obtained using the fact that $\boldsymbol{\beta}$ is Lipschitz continuous to infer $\|\boldsymbol{\beta}-\pi_T^0\boldsymbol{\beta}\|_{L^\infty(T)^d}\leqslant L_{\boldsymbol{\beta},T}h_T$ followed by the inverse inequality (6) together with the definition (23) of $\tau_{\mathrm{ref},T}$.

To treat $\mathfrak{T}_{2,2}$ and $\mathfrak{T}_{2,3}$, we proceed differently according to the value of the local Péclet number. We write $\mathfrak{T}_{2,2}=\mathfrak{T}_{2,2}^{\mathrm{d}}+\mathfrak{T}_{2,2}^{\mathrm{a}}$ and $\mathfrak{T}_{2,3}=\mathfrak{T}_{2,3}^{\mathrm{d}}+\mathfrak{T}_{2,3}^{\mathrm{a}}$, where the superscript "a" corresponds to integrals where $|\mathrm{Pe}_{TF}|\leqslant1$, while the superscript "d" corresponds to integrals where $|\mathrm{Pe}_{TF}|>1$ (which conventionally include all faces where $\nu_F=0$). We denote by $\mathbf{1}_{|\mathrm{Pe}_{TF}|\leqslant1}$ and $\mathbf{1}_{|\mathrm{Pe}_{TF}|>1}$ the two characteristic functions of these regions. The idea is that we use the diffusive norm of $\underline{\mathsf{v}}_h$ if $|\mathrm{Pe}_{TF}|\leqslant1$, whereas we use the advective norm if $|\mathrm{Pe}_{TF}|>1$. Before proceeding, we observe that, for all $T\in\mathcal{T}_h$ and all $F\in\mathcal{F}_h$, the following holds:

$$\|\widehat{\mathsf{u}}_F-\widehat{\mathsf{u}}_T\|_F=\|\pi_F^k(u_{|T_\nu(F)}-\widehat{\mathsf{u}}_T)\|_F\leqslant\|u_{|T_\nu(F)}-\widehat{\mathsf{u}}_T\|_F,\tag{61}$$

where we have used that $\widehat{\mathsf{u}}_F=\pi_F^k u_{|T_\nu(F)}$ (see (28)) $\widehat{\mathsf{u}}_{T|F}\in\mathbb{P}_{d-1}^k(F)$, and that π_F^k is a projector. For $\mathfrak{T}_{2,2}^{\mathrm{d}}$, it is also useful to notice that, since $A^-(0)=0$,

$$\left|\tfrac{\nu_F}{h_F}A^-(\mathrm{Pe}_{TF})\right|\lesssim\tfrac{\nu_F}{h_F}|\mathrm{Pe}_{TF}|=|\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF}|\leqslant\beta_{\mathrm{ref},T}\tag{62}$$

whenever $\nu_F > 0$ (which is always the case if $\mathbf{1}_{|\mathrm{Pe}_{TF}|\leqslant 1} \not\equiv 0$). Hence, observing that $\nu_F > 0$ indicates that the exact solution u does not jump across F, so that we can simply write $u_{|T}$ in place of $u_{|T_\nu(F)}$,

$$
\begin{aligned}
&|\mathfrak{T}_{2,2}^{\mathrm{d}}| + |\mathfrak{T}_{2,3}^{\mathrm{d}}| \\
&\quad\leqslant \sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} (|\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF}|\mathbf{1}_{|\mathrm{Pe}_{TF}|\leqslant 1}\,|u_{|T} - \widehat{\mathsf{u}}_T|, |\mathsf{v}_F - \mathsf{v}_T|)_F \\
&\qquad + \mathrm{Lip}(A^-)\sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} \tfrac{\nu_F}{h_F}(|\mathrm{Pe}_{TF}|\mathbf{1}_{|\mathrm{Pe}_{TF}|\leqslant 1}\,|\widehat{\mathsf{u}}_F - \widehat{\mathsf{u}}_T|, |\mathsf{v}_F - \mathsf{v}_T|)_F \\
&\quad\lesssim \left\{\sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} \tfrac{\nu_F}{h_F}\|\mathrm{Pe}_{TF}\mathbf{1}_{|\mathrm{Pe}_{TF}|\leqslant 1}\|_{L^\infty(F)}^2\|u_{|T} - \widehat{\mathsf{u}}_T\|_F^2\right\}^{1/2} \times \left\{\sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} \tfrac{\nu_F}{h_F}\|\mathsf{v}_F - \mathsf{v}_T\|_F^2\right\}^{1/2} \\
&\quad\lesssim \left\{\sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} \beta_{\mathrm{ref},T}\min(1,\mathrm{Pe}_T)\|u_{|T} - \widehat{\mathsf{u}}_T\|_F^2\right\}^{1/2} \|\underline{\mathsf{v}}_h\|_{\nu,h},
\end{aligned}
\tag{63}
$$

where we have used (61) and (62) to bound the second addend in the first line and the norm equivalence (21) to conclude. To estimate $\mathfrak{T}_{2,2}^{\mathrm{a}}$, it suffices to observe that

$$
\begin{aligned}
|\mathfrak{T}_{2,2}^{\mathrm{a}}| &\lesssim \left\{\sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} \|\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF}\mathbf{1}_{|\mathrm{Pe}_{TF}|>1}\|_{L^\infty(F)}\|u_{|T} - \widehat{\mathsf{u}}_T\|_F^2\right\}^{1/2} \\
&\qquad\times \left\{\sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} \||\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF}\mathbf{1}_{|\mathrm{Pe}_{TF}|>1}|^{1/2}(\mathsf{v}_F - \mathsf{v}_T)\|_F^2\right\}^{1/2} \\
&\lesssim \left\{\sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} \beta_{\mathrm{ref},T}\min(1,\mathrm{Pe}_T)\|u_{|T} - \widehat{\mathsf{u}}_T\|_F^2\right\}^{1/2} \|\underline{\mathsf{v}}_h\|_{\boldsymbol{\beta},\mu,h},
\end{aligned}
\tag{64}
$$

where the introduction of the advective norm in the last inequality is justified since, owing to (38) (see also (14) if $\nu_F = 0$) and Assumption (A2),

$$
|\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF}|\mathbf{1}_{|\mathrm{Pe}_{TF}|>1} \lesssim \tfrac{\nu_F}{h_F}|A|(\mathrm{Pe}_{TF}). \tag{65}
$$

To estimate $\mathfrak{T}_{2,3}^{\mathrm{a}}$, recalling (61), we observe that

$$
|\mathfrak{T}_{2,3}^{\mathrm{a}}| \leqslant \sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} (|\tfrac{\nu_F}{h_F}A^-(\mathrm{Pe}_{TF})|\mathbf{1}_{|\mathrm{Pe}_{TF}|>1}|\pi_F^k(u_{|T_\nu(F)} - \widehat{\mathsf{u}}_T)|, |\mathsf{v}_F - \mathsf{v}_T|)_F.
$$

For given $T \in \mathcal{T}_h$ and $F \in \mathcal{F}_T$, we have the following mutually exclusive cases: (i) $\nu_F > 0$ or ($\nu_F = 0$ and $F \subset \mathcal{I}_{\nu,\boldsymbol{\beta}}^+$), in which case $u_{|T_\nu(F)} = u_{|T}$ since u does not have a jump at F (see (2c) if $F \subset \mathcal{I}_{\nu,\boldsymbol{\beta}}^+$); (ii) $\nu_F = 0$ and $F \subset \mathcal{I}_{\nu,\boldsymbol{\beta}}^-$, in which case, recalling (14), $\frac{\nu_F}{h_F}A^-(\mathrm{Pe}_{TF}) = (\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF})^- = 0$. Hence, in any case, $|\frac{\nu_F}{h_F}A^-(\mathrm{Pe}_{TF})||\pi_F^k(u_{|T_\nu(F)} - \widehat{\mathsf{u}}_T)| = |\frac{\nu_F}{h_F}A^-(\mathrm{Pe}_{TF})||\pi_F^k(u_{|T} - \widehat{\mathsf{u}}_T)|$. Using this fact, and observing that, for all $T \in \mathcal{T}_h$ and all $F \in \mathcal{F}_T$, $|\frac{\nu_F}{h_F}A^-(\mathrm{Pe}_{TF})| \lesssim |\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF}| \lesssim \beta_{\mathrm{ref},T}$, we infer the estimate

$$
|\mathfrak{T}_{2,3}^{\mathrm{a}}| \lesssim \left\{\sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} \beta_{\mathrm{ref},T}\min(1,\mathrm{Pe}_T)\|u_{|T} - \widehat{\mathsf{u}}_T\|_F^2\right\}^{1/2} \|\underline{\mathsf{v}}_h\|_{\boldsymbol{\beta},\mu,h}. \tag{66}
$$

To conclude the estimate on $\mathfrak{T}_{2,2}$, we collect the bounds (63), (64), and (66), and invoke (7) to write $\|u_{|T} - \widehat{\mathsf{u}}_T\|_F \leqslant C_{\mathrm{app}}h_T^{k+1/2}|u|_{H^{k+1}(T)}$, so that

$$
|\mathfrak{T}_{2,2}| + |\mathfrak{T}_{2,3}| \lesssim \left\{\sum_{T\in\mathcal{T}_h} \beta_{\mathrm{ref},T}\min(1,\mathrm{Pe}_T)h_T^{2k+1}\|u\|_{H^{k+1}(T)}^2\right\}^{1/2} \|\underline{\mathsf{v}}_h\|_{\sharp,h}. \tag{67}
$$

(iii) *Weakly enforced boundary conditions.* Let us now estimate $\mathfrak{T}_3$. Denoting by $\mathfrak{T}_{3,1}$ and $\mathfrak{T}_{3,2}$ the two addends in $\mathfrak{T}_3$, the estimate of $\mathfrak{T}_{3,1}$ is a straightforward consequence of the Cauchy-Schwarz inequality, the definition (20) $\|\cdot\|_{\nu,h}$, and the approximation property (9) of $\breve{u}_T = p_T^k \widehat{\underline{u}}$:

$$|\mathfrak{T}_{3,1}| \leqslant \left\{ \sum_{F\in\mathcal{F}_h^{\rm b}} \nu_F h_F \|\boldsymbol{\nabla}(u - \breve{u}_{T(F)})\|_F^2 \right\}^{1/2} \|\underline{\mathsf{v}}_h\|_{\nu,h} \leqslant \left\{ \sum_{F\in\mathcal{F}_h^{\rm b}} \nu_F h_{T(F)}^{2(k+2)} \|u\|_{H^{k+2}(T(F))}^2 \right\}^{1/2} \|\underline{\mathsf{v}}_h\|_{\sharp,h}. \tag{68}$$

To estimate $\mathfrak{T}_{3,2}$, we apply similar ideas to those employed to bound $\mathfrak{T}_{2,2}$. We first observe that, for all $F \in \mathcal{F}_h^{\rm b}$,

$$\|u - \widehat{\mathsf{u}}_F\|_F \lesssim h_T^{k+1/2} |u|_{H^{k+1}(T)}. \tag{69}$$

Since $|\frac{\nu_F}{h_F} A^+(\mathrm{Pe}_{TF})| \lesssim |\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF}|$ (proved as for A^- above) and $|A^+(\mathrm{Pe}_{TF})| \lesssim |\mathrm{Pe}_{TF}|$ whenever $\nu_F > 0$, invoking the definitions (20) and (22) of the diffusive and advective norms and reasoning as in the estimates of $\mathfrak{T}_{2,2}^{\rm d}$ and $\mathfrak{T}_{2,2}^{\rm a}$, estimate (65) and the approximation property (69) yield

$$\begin{aligned}
|\mathfrak{T}_{3,2}| &\lesssim \sum_{F\in\mathcal{F}_h^{\rm b}} \tfrac{\nu_F}{h_F} (|\mathrm{Pe}_F| \mathbf{1}_{|\mathrm{Pe}_F|\leqslant 1} |\widehat{\mathsf{u}}_F - u|, |\mathsf{v}_F|)_F + \sum_{F\in\mathcal{F}_h^{\rm b}} (|\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF}| \mathbf{1}_{|\mathrm{Pe}_F|>1} |\widehat{\mathsf{u}}_F - u|, |\mathsf{v}_F|)_F \\
&\lesssim \left\{ \sum_{F\in\mathcal{F}_h^{\rm b}} \beta_{{\rm ref},T} \|\mathrm{Pe}_F \mathbf{1}_{|\mathrm{Pe}_F|\leqslant 1}\|_{L^\infty(F)} h_T^{2k+1} \|u\|_{H^{k+1}(T)}^2 \right\}^{1/2} \|\underline{\mathsf{v}}_h\|_{\nu,h} \\
&\quad + \left\{ \sum_{F\in\mathcal{F}_h^{\rm b}} \|\boldsymbol{\beta}\cdot\boldsymbol{n}_{TF} \mathbf{1}_{|\mathrm{Pe}_F|>1}\|_{L^\infty(F)} h_T^{2k+1} \|u\|_{H^{k+1}(T)}^2 \right\}^{1/2} \|\underline{\mathsf{v}}_h\|_{\boldsymbol{\beta},\mu,h} \\
&\lesssim \left\{ \sum_{F\in\mathcal{F}_h^{\rm b},\, F\subset\partial T} \beta_{{\rm ref},T} \min(1, \mathrm{Pe}_T) h_T^{2k+1} \|u\|_{H^{k+1}(T)}^2 \right\}^{1/2} \|\underline{\mathsf{v}}_h\|_{\sharp,h}.
\end{aligned} \tag{70}$$

The proof is completed by plugging estimates (59), (60), (67), (68), and (70) into (58), and using the resulting bound to estimate the right-hand side of (57).

References

[1] F. Bassi, S. Rebay, G. Mariotti, S. Pedinotti, and M. Savini. A high-order accurate discontinuous finite element method for inviscid and viscous turbomachinery flows. In R. Decuypere and G. Dibelius, editors, *Proceedings of the* 2nd *European Conference on Turbomachinery Fluid Dynamics and Thermodynamics*, pages 99–109, 1997.

[2] L. Beirão da Veiga, J. Droniou, and M. Manzini. A unified approach for handling convection terms in finite volumes and mimetic discretization methods for elliptic problems. *IMA J. Numer. Anal.*, 31(4):1357–1401, 2010.

[3] F. Brezzi, K. Lipnikov, and M. Shashkov. Convergence of the mimetic finite difference method for diffusion problems on polyhedral meshes. *SIAM J. Numer. Anal.*, 43(5):1872–1896, 2005.

[4] C. Chainais-Hillairet and J. Droniou. Finite-volume schemes for noncoercive elliptic problems with Neumann boundary conditions. *IMA J. Numer. Anal.*, 31(1):61–85, 2011.

[5] Y. Chen and B. Cockburn. Analysis of variable-degree HDG methods for convection-diffusion equations. part II: semimatching nonconforming meshes. *Math. Comp.*, 83(285):87–111, 2014.

[6] B. Cockburn, B. Dong, J. Guzmán, M. Restelli, and R. Sacco. A hybridizable discontinuous Galerkin method for steady-state convection-diffusion-reaction problems. *SIAM J. Sci. Comput.*, 31(5):3827–3846, 2009.

[7] D. A. Di Pietro and A. Ern. *Mathematical aspects of discontinuous Galerkin methods*, volume 69 of *Mathématiques & Applications*. Springer-Verlag, Berlin, 2012.

[8] D. A. Di Pietro and A. Ern. Hybrid high-order methods for variable-diffusion problems on general meshes. *C. R. Acad. Sci Paris, Ser. I*, 2014. To appear. DOI 10.1016/j.crma.2014.10.013.

[9] D. A. Di Pietro and A. Ern. A hybrid high-order locking-free method for linear elasticity on general meshes. *Comput. Meth. Appl. Mech. Engrg.*, 283:1–21, 2015.

[10] D. A. Di Pietro, A. Ern, and J.-L. Guermond. Discontinuous Galerkin methods for anisotropic semi-definite diffusion with advection. *SIAM J. Numer. Anal.*, 46(2):805–831, 2008.

[11] D. A. Di Pietro, A. Ern, and S. Lemaire. An arbitrary-order and compact-stencil discretization of diffusion on general meshes based on local reconstruction operators. *Comput. Meth. Appl. Math.*, 14(4):461–472, 2014.

[12] D. A. Di Pietro and S. Lemaire. An extension of the Crouzeix–Raviart space to general meshes with application to quasi-incompressible linear elasticity and Stokes flow. *Math. Comp.*, 84(291):1–31, 2015.

[13] J. Droniou. Remarks on discretizations of convection terms in hybrid mimetic mixed methods. *Netw. Heterog. Media*, 5(3):545–563, 2010.

[14] J. Droniou and R. Eymard. A mixed finite volume scheme for anisotropic diffusion problems on any grid. *Numer. Math.*, 105:35–71, 2006.

[15] J. Droniou, R. Eymard, T. Gallouët, and R. Herbin. A unified approach to mimetic finite difference, hybrid finite volume and mixed finite volume methods. *M3AS Mathematical Models and Methods in Applied Sciences*, 20(2):1–31, 2010.

[16] T. Dupont and R. Scott. Polynomial approximation of functions in Sobolev spaces. *Math. Comp.*, 34(150):441–463, 1980.

[17] A. Ern and J. Proft. Multi-algorithmic methods for coupled hyperbolic-parabolic problems. *Int. J. Numer. Anal. Model.*, 1(3):94–114, 2006.

[18] R. Eymard, T. Gallouët, and R. Herbin. Discretization of heterogeneous and anisotropic diffusion problems on general nonconforming meshes. SUSHI: a scheme using stabilization and hybrid interfaces. *IMA J. Numer. Anal.*, 30(4):1009–1043, 2010.

[19] F. Gastaldi and A. Quarteroni. On the coupling of hyperbolic and parabolic systems: Analytical and numerical approach. *Appl. Numer. Math.*, 6:3–31, 1989.

[20] R. Herbin and F. Hubert. Benchmark on discretization schemes for anisotropic diffusion problems on general grids. In R. Eymard and J.-M. Hérard, editors, *Finite Volumes for Complex Applications V*, pages 659–692. John Wiley & Sons, 2008.

[21] P. Houston, C. Schwab, and E. Süli. Discontinuous *hp*-finite element methods for advection-diffusion-reaction problems. *SIAM J. Numer. Anal.*, 39(6):2133–2163, 2002.

[22] C. Johnson and J. Pitkäranta. An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation. *Math. Comp.*, 46(173):1–26, 1986.

[23] J. Wang and X. Ye. A weak Galerkin element method for second-order elliptic problems. *J. Comput. Appl. Math.*, 241:103–115, 2013.